APPLICATION OF A LASER SCANNER TO THREE DIMENSIONAL VISUAL SENSING TASKS

NAGw-1333

by

Arthur M. Ryan

Rensselaer Polytechnic Institute
Electrical, Computer, and Systems Engineering Department
Troy, New York 12180-3590

May 1992

CIRSSE REPORT #112

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENT

"No man is an island entire of itself" - John Donne

I wish to express my thanks to...

The faculty, students and staff of CIRSSE for their comments and assistance during the preparation of this research,

Dr. Lester Gerhardt for his guidance and advice,

And most of all to my wife, Melinda. I shall not compare thee to a summer's day, for thou art *far* more lovely and more temperate.

ABSTRACT

This report describes the issues associated with using a laser scanner for visual sensing and the methods developed by the author to address them. A laser scanner is a device that controls the direction of a laser beam by deflecting it through a pair of orthogonal mirrors, the orientations of which are specified by a computer. If a calibrated laser scanner is combined with a calibrated camera, it is possible to perform three dimensional sensing by directing the laser at objects within the field of view of the camera. There are several issues associated with using a laser scanner for three dimensional visual sensing that must be addressed in order to use the laser scanner effectively. First, methods are needed to calibrate the laser scanner and estimate three dimensional points. Second, methods to estimate three dimensional points using a calibrated camera and laser scanner are required. Third, methods are required for locating the laser spot in a cluttered image. Fourth, mathematical models that predict the laser scanner's performance and provide structure for three dimensional data points are necessary. The author has developed several methods to address each of these and has evaluated these methods to determine how and when they should be applied. The theoretical development, implementation, and results when used in a dual arm eighteen degree of freedom robotic system for space assembly is described.

1. Introduction

In performing robotic assembly tasks, it is necessary to provide a reliable three dimensional sensing capability to produce accurate information about the locations of objects or features within the workspace of a robot. At the Center for Intelligent Robotic Systems for Space Exploration (CIRSSE), research has been directed toward creating a three dimensional visual sensing system for robotics assembly. Currently, this system collects information using passive sensing techniques employing multiple cameras. The visual sensing system is also equipped with a laser scanner that provides the ability to generate and introduce structured light into the observed scene. A laser scanner is a device that controls the direction of a laser beam by deflecting it through a pair of orthogonal mirrors, the orientations of which can be specified by a computer. A laser scanner operated in conjunction with a camera can provide an active sensing capability that can complement a purely passive multiple camera system. However, to effectively use a laser scanner in three dimensional visual sensing, several problems first need to be solved.

The research presented in this report describes the issues involved in applying a laser scanner to active three dimensional visual sensing and the methods developed to address them. The description of this research is divided into six chapters. The first chapter surveys the current state of the art in laser scanner technology and summarizes the early laser scanner research conducted at CIRSSE that provided the motivation for the current research. This chapter also defines the objectives and scope of research presented in this report, and describes the facilities used for this research. The second chapter examines the mathematical relationships that govern the operation of a laser scanner and how these relationships can be used to calibrate the laser scanner. The third chapter describes the engineering problems

involved in using a laser scanner for three dimensional sensing, presents methods for addressing these problems, and evaluates these methods to determine their merits and limitations. The fourth chapter presents two sensing applications that employ the laser scanner and illustrates how the methods developed in the second and third chapters can be applied to these applications. The fifth chapter evaluates the laser scanner as a sensing tool in the context of alternative passive and active visual sensing techniques. The last chapter briefly summarizes the results of this research and presents possible opportunities for future work in this area.

1.1 History of Laser Scanner Technology and Research

Laser scanning entails deflecting a laser beam in a well defined and controlled manner. A laser scanner can direct the beam in one or two dimensions. There are a variety of approaches used to implement a laser scanner, and these are discussed in this section. The methods can be broken down into three major types: rotary scanners, two axis reflective plate, and dual orthogonal scanning mirrors. One of these scanning implementations was chosen for the CIRSSE testbed and the reasons for this selection will be described. Finally a brief summary of the early laser scanner research performed at CIRSSE is presented to provide the motivation for the research presented in this report.

One method used for laser scanning is a rotary scanning mechanism. The scanning mechanism can consist of a polygonal mirror [1] that rotates or oscillates and thereby deflects an incident laser beam in one axis. This deflected beam can then be deflected in a second axis using a monogonal mirror (i.e. a simple flat mirror)[2]. This approach permits rapid scanning in one axis and slow scanning in the second axis, much like the raster scan of a video signal. These scanning mechanisms are used in image scanners and laser typesetters and printers. These

scanners are highly sensitive to vibration or improper balancing of the mirror either of which causes the rotating mirror to wobble about its rotational axis. This wobble cause the scanned laser beam to assume a curved path rather than a straight line. Indeed, eliminating this wobble either through improved manufacturing processes or optical assemblies that can correct for it[3] is a current research topic.

An alternate approach to rotary scanning is to pass the laser beam though a rotating holographic plate (this approach is called hologonic scanning). The plate is etched in such a way as to deflect the laser scanner in different directions as the plate rotates. This device is subject to rotational wobble for the same reasons as with the polygonal mirror, but because the hologonic scanner is a "transmissive" optical device [4] as opposed to reflective, the hologonic scanner displays less error than the polygonal mirror. Rotary scanners are extremely effective for high speed high accuracy scanning, but they are not designed to direct the laser beam at a single point and maintain its position.

Another laser scanning method recently developed uses a reflective plate which can be rotated in two axes [5] [6]. The plate is moved using magnets or electro-mechanical actuators. Each axis is equipped with a positional sensor to permit the plate to be positioned to specific orientations. These devices exhibit high resolution (11μdegrees) and are compact and simple. The main drawback of this configuration is that the range of motion of the scanning plates is currently between ± 1.48 and ± 3 degrees.

The third laser scanning method employs two planar mirrors rotated by galvanometers. A galvanometer contains an electro-mechanical actuator and a position transducer, thereby permitting the mirror to be rotated to specific positions with a high resolution (typically .00977 degrees). The two mirrors are positioned to be orthonormal to each other. A laser beam is deflected by the first mirror and then

passes to the second mirror where it is again deflected before leaving the scanning assembly. Essentially, each mirror deflects the laser beam in one axis. There are minor variations on the orthogonal mirror approach such as placement of mirrors and lenses [7] [8], but all of the dual axis mirror approaches have one thing in common: they can deflect the laser beam over a wide angular arc (typically $\pm$10degrees per axis). This characteristic of the two mirror laser scanner combined with its ability to orient the laser beam in specific directions (via the use of galvanometers) makes it ideal for visual sensing applications since it can project the laser beam over a large workspace in a well defined and controlled manner. The specific dynamics of the dual mirror laser scanner will be discussed in greater detail in chapter 2.

Research at CIRSSE involving the laser scanner began shortly after the vision system was installed in the CIRSSE testbed during 1990. Initially, the laser scanner was used to generate discrete or continuous patterns such as grid lines or predefined shapes. This pattern generating capability is possible since the laser scanner is equipped with a shutter that can block the laser beam when blanking is required between points. At this early phase of the research, the laser could be controlled only in terms of its scanning mirrors. Initially, there was no ability to control the laser in terms of cartesian space.

The next phase of the research, initiated by the author, was directed towards using a camera as a feedback mechanism to direct the laser at a specific pixel coordinate(valuable for visual servoing among other uses). The laser was directed in the field of view of a camera and the camera was used to identify the reflection of the laser beam off objects in the workspace (this reflection is colloquially referred to as the *laser spot*). Control of the laser was still in terms of the angles of the scanning mirrors and these angles were repeatedly adjusted by the computer until the laser spot was centered onto the desired pixel coordinates. The process of locating the

laser spot in the camera image was premised on the assumption that the laser spot was the brightest object in the image. Hence, this laser and camera configuration operated best under subdued lighting conditions. This work highlighted several issues that served as the basis for the current laser research:

- The CIRSSE vision system was capable of using the laser and cameras in a well coordinated manner.
- Using a camera to detect the laser spot can be a valuable sensing capability since the laser could illuminate objects that the camera might otherwise be unable to distinguish.
- The current camera feedback method assumed that the laser spot is the brightest object in the image. Methods must be developed that circumvent this assumption to permit a laser and camera to operate in a wider variety of environmental conditions.
- If the laser could be calibrated to the same coordinate system as the camera, it would then be possible to directly place the laser at a world point and use a calibrated camera to confirm the proper placement of the laser. Further, because the laser would be calibrated, it would be unnecessary to repeatedly direct the laser to settle on a desired point.

With these ideas in mind, subsequent research was directed toward using the laser scanner for active three dimensional visual sensing. The specific goal of this research and a description of the approach used to achieve it are described in the next section.

1.2 Objective of Laser Scanner Research

Three dimensional visual sensing methods can be broken down into two major categories: passive techniques and active techniques. A passive technique is one that uses available light sources (i.e. general illumination) while an active technique employs the projection of some externally supplied prestructured light [9]. The passive visual sensing technique employed at CIRSSE uses two or more cameras oriented in such a way as to permit all the cameras to view a common area of the workspace. Features identified in one camera image are corresponded to similar features in the images of the other cameras. Stereoscopic triangulation is then employed to identify the three dimensional coordinates of the object corresponding to the feature identified in the images. This technique works well as long as the features can be correctly identified in the camera images. Identification, however, can be affected by adverse lighting conditions or by occlusion. In such cases, active sensing techniques can be employed to clarify the situation by injecting a well defined light signature into the workspace and using the camera to detect the presence or distortion of this signature.

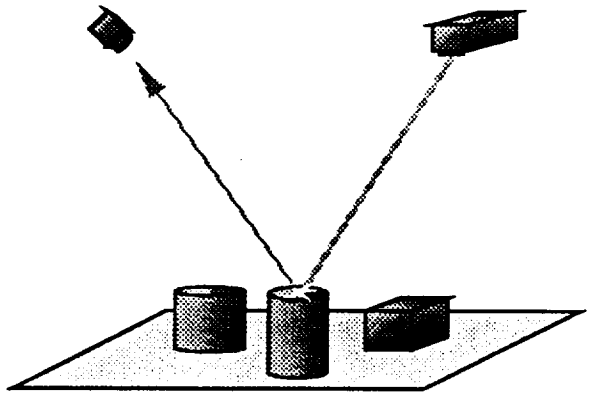


Figure 1.1: Concept of active visual sensing using a laser scanner

The research presented in this report focuses on using the laser scanner in

conjunction with a camera to perform active three dimensional visual sensing. The concept of this application is presented in figure 1.1. The laser scanner directs the laser beam in a well defined manner and a camera is used to detect the reflection of the laser beam as it strikes objects in the workspace (this reflection is referred to as the *laser spot*). If it is possible to ascertain the directions of both the laser beam and the projected ray from the camera's image plane to the laser spot, then is is possible to determine the three dimensional coordinates of the laser spot. Since the laser spot is generated by the reflection of the laser beam off an object, it can be assumed that there is some feature at the coordinates of the laser spot. If the laser beam is directed to many different places in the workspace, it is possible to obtain three dimensional information about a region of the workspace.

While the concept of using the laser scanner for active sensing is straightforward, successfully realizing this concept involves solving a variety of different problems. The relevant issues and their significance are enumerated below:

1. Identify the mathematical relationships that govern the operation of the laser scanner. Accomplishing this task will provide the means to control the direction of the laser beam.

2. Calibrate the laser scanner. This will provide the means of mathematically describing the origin and direction of the laser beam given a specific orientation of the scanning mirrors.

3. Develop methods to estimate three dimensional points using a calibrated laser and camera. These methods will provide the means for identifying the three dimensional coordinates of the laser spot.

4. Develop computer simulation models of the laser scanner and the laser / camera sensing configuration. These models will provide a means to predict the

performance of the laser scanner and the reliability of the point estimates generated by a laser and camera.

5. Develop methods to identify the laser spot in a camera's image in the presence of extraneous noise. These techniques will enhance the reliability of the laser / camera configuration by making it more robust to unexpected scene and lighting conditions.

6. Identify data structures and models for representing three dimensional information. Such structures can enhance the three dimensional information provided by a set of point estimates generated by the laser and camera.

Successfully addressing these issues will provide an active three dimensional sensing capability that is both useful and well understood. Some of the methods used to solve these problems draw from techniques used in multiple camera passive sensing. This is appropriate since a camera maps three dimensional points to two dimensional pixel coordinates, and similarly, a laser scanner maps three dimensional points to two scanning mirror angles. It is not surprising, therefore, that some of the passive camera techniques can be adapted to a laser scanner. However, some methods have to be developed from scratch either because problems associated with a laser scanner are different from those of cameras, or new methods can take advantage of some unique properties of the laser scanner. During the course of this report, techniques that are based on passive camera methodology will be distinguished from techniques that were developed specifically for the laser scanner.

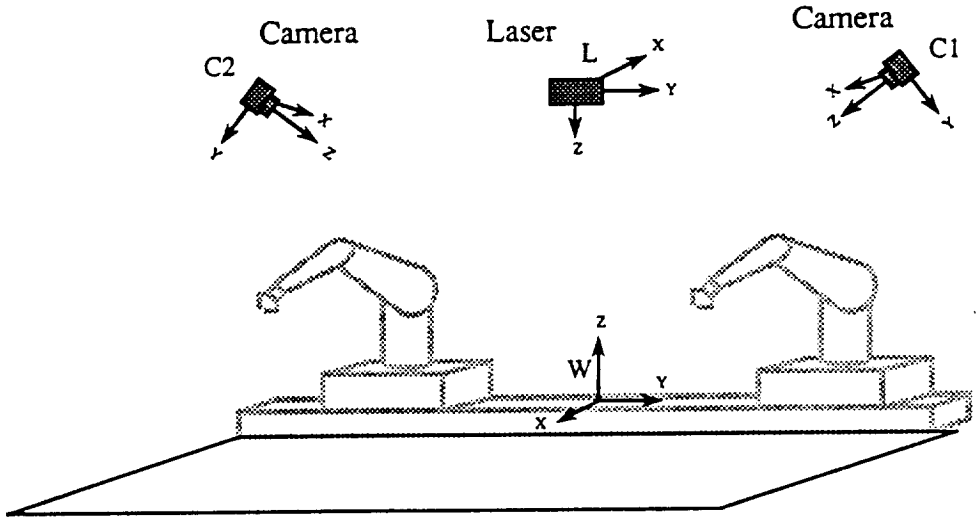


Figure 1.2: CIRSSE testbed

1.3 Description of the CIRSSE Testbed

The CIRSSE testbed is designed to support research in robotic assembly tasks for space applications. The major components of the testbed are depicted in figure 1.2. The centerpiece of the testbed is a pair of PUMA robots each of which is mounted on a movable cart. Both carts are mounted on a twelve foot track thereby permitting the robots to operate over a large work volume. Five cameras are positioned throughout the testbed. Two cameras are located on the ceiling, two on the wrist of one of the PUMA robots and one on the second PUMA robot. A laser scanner, manufactured by General Scanning[7] is mounted on the ceiling in between the two ceiling cameras The two ceiling cameras and the laser scanner are represented in the figure. Color plates A.1 and A.2 in appendix A also show the physical configuration of the equipment in the CIRSSE testbed.

The coordinate systems relevant to the laser scanner research are denoted in figure 1.2 (L, W, $C1$, $C2$). For the purposes of system calibration and three

dimensional point estimation, the CIRSSE testbed has a well defined world origin point W. This point is located at the center of the track utilized by the robot carts. The exact location and orientation of this point is described in a CIRSSE technical memorandum[10]. This report will refer to the testbed's world point as either the *world origin* or the *world coordinate system.*

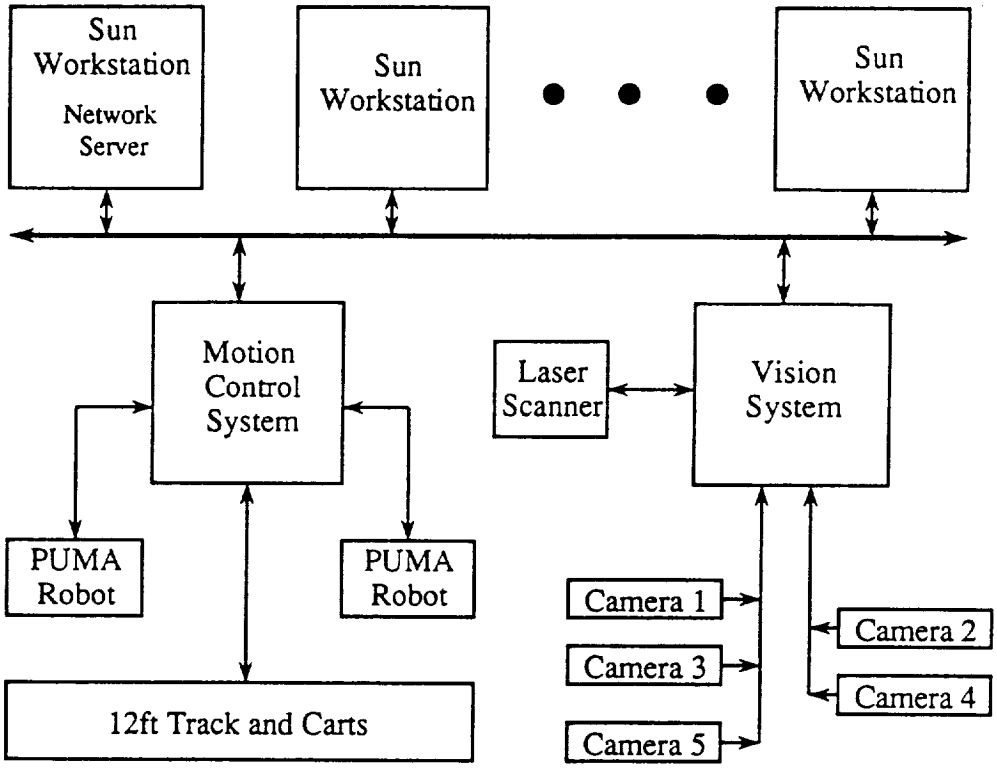


Figure 1.3: CIRSSE computer network

CIRSSE has a distributed computer system of Sun workstations and 68xxx series processors. The computer system is shown in figure 1.3. The cameras and laser scanner are controlled by a vision system that contains two 68xxx series processors and a suite of image processing boards manufactured by Datacube. The robots are controlled by the motion control system consisting of five 68xxx series processors and a variety of I/O interface boards. The vision and motion control systems are

connected via ethernet to the Sun workstations. Communication between the different components of the computer system is performed using the CIRSSE Testbed Operating System (CTOS) which operates in conjunction with UNIX on the Sun workstations and VxWorks on the motion control and vision systems.

One of the objectives of CIRSSE's research in robotics assembly is the creation of a hierarchically intelligent machine in which different types of decisions are carried out at different strata of processing. At the highest level is a representation and planning level which determines the necessary operations to perform a specific assembly task and a plan as to how these steps should be executed. This plan is then passed down to a coordination level which executes the plan by orchestrating vision and motion operations. Finally, below the coordinator is the execution level which carries out the low level operations of robot motion, laser control and image acquisition / processing to carry out the assembly task. The laser scanner research resides primarily at the execution level, although the computer simulation models are more directed at the coordination and representation levels.

2. Calibration of a Laser Scanner

A laser is useful in 3-D visual sensing because it provides an active sensing capability. The laser emits a beam of light that a camera can detect as it reflects off objects within the camera's field of view. An active sensing configuration, such as a laser and camera, can enhance the reliability and flexibility of a vision system since it can generate structured light and "ground truth". In some applications, the laser does not have to be calibrated[9], but calibration is necessary to fully utilize the capabilities of the laser.

The phrase "calibrated laser" is misleading in that the laser itself is not calibrated. It is usually incorporated in an assembly that can direct the laser beam in some well-defined manner. When the laser is calibrated, the entire assembly is actually calibrated. One useful laser assembly is a laser scanner[7] which uses mirrors to deflect the laser beam in a controlled manner.

Calibrating a laser scanner is similar to calibrating a camera in that both devices have intrinsic and extrinsic parameters. Calibration parameters can be broken into two different classes: intrinsic and extrinsic. Intrinsic parameters are inherent in the design of a device and do not change unless the internal configuration of the device is physically altered. Extrinsic parameters describe the relationship between the device and its surrounding environment and these change whenever this relationship changes. For a laser scanner, intrinsic parameters include the distance between the scanner's mirrors and the relationships between the mirror's rotation and the voltage applied to their rotational mechanisms (galvanometers). A laser scanner's extrinsic parameters describe the pose of the laser scanner with respect to some coordinate frame. To better understand the process of laser scanner calibration, it

is instructive to examine the internal arrangement of the device and identify the mathematical relationships that describe its operation.

2.1 The Mathematical Model of a Laser Scanner

The internal arrangement of the laser scanner is depicted in Figure 2.1. The device consists of a laser and two scanning mirrors. Each mirror is connected to a galvanometer that rotates the mirror as a function of a control voltage applied to it. The mirrors are configured such that their rotational axes are mutually orthogonal. The rotational axis of the θ_x mirror is parallel with the z axis of the laser scanner, and the rotational axis of the θ_y mirror is parallel to the x axis of the laser scanner.

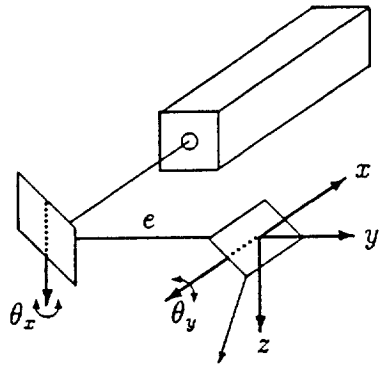
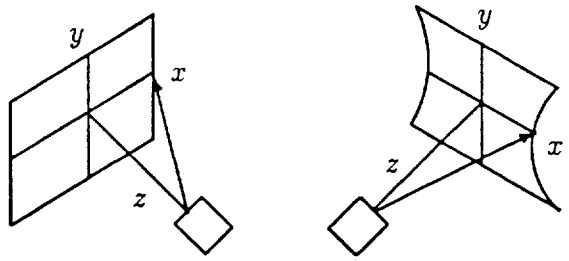


Figure 2.1: Internal arrangement of a laser scanner

Prior research with laser scanners, [11], highlighted the problems associated with creating mathematical relationships between the deflection of the scanning mirrors and the vector of the outgoing laser ray in closed form. If the laser and mirrors are placed at arbitrary locations and orientations, these mathematical relationships become intractable. This is due, in part, to the difficulty in determining the values of some of the necessary parameters[11]. To alleviate these problems, two

constraints on the placement of the laser and mirrors must be established. First, the beam emitted by the laser must be parallel to the rotational axis of the θ_y mirror. Second, the laser beam must intersect each mirror at a point along its rotational axis. These constraints are reasonable and practical when one considers that a laser scanner can be assembled with high precision using current manufacturing technology. These constraints reduce the mathematical relationships between the scanning mirrors and the outgoing laser ray to two simple relationships which are described later in more detail.



With Pincushion Compensation Without Pincushion Compensation

Figure 2.2: Effect of pincushion distortion

When the laser beam is scanned onto a plane normal to the laser's z axis and at some fixed distance from the laser's origin, it is possible to determine the (x, y) coordinates of the laser spot as a function of the mirror's angular deflection. In this situation, the y coordinate of the laser spot is strictly a function of θ_y and the distance to the plane as shown below:

$$y = z \tan \theta_y \tag{2.1}$$

where z is the distance from the laser's origin (L) to the plane normal to the laser's z axis.

If the x coordinate of the spot is assumed to be independent of θ_y, then the laser exhibits *pincushion distortion* (as depicted in Figure 2.2). In reality, the x

coordinate of the spot is a function of both θ_x and θ_y. This interdependency is due to the fact that the laser beam strikes the θ_y mirror after it is deflected by the θ_x mirror (see Figure 2.1). The expression for the x coordinate of the laser spot is

$$x = (z \sec \theta_y + e) \tan \theta_x \tag{2.2}$$

where z is the distance from L as in (2.1), and e is the distance between the two scanning mirrors. Note that the $z \sec \theta_y$ term increases as $|\theta_y|$ increases. Hence, the displacement of the laser spot "flares out" away from the origin in the x direction as the spot moves away from the origin in the y direction. Equations (2.1) and (2.2) provide the necessary relationships to direct the laser spot to any 3-D point defined with respect to the laser coordinate frame without pincushion distortion.

The mathematical relationships presented above describe the direction of the laser ray. Hence, while it is possible to determine the value of the scanning mirrors angles given a three dimensional point, the mathematical relationships are not closed since a given set of mirror angles does not relate to a unique three dimensional point. This is an inherent property of the laser scanner since the mirrors control the direction of the laser beam, and there are an infinite number of three dimensional points that are colinear with this beam. Closed form solutions are only possible if one of the axes (usually the Z axis) of the three dimensional point is fixed.

2.2 Calibration of the Intrinsic Parameters

The distance between the two mirrors (e) can be obtained by direct measurement. The degree to which errors in this measurement will affect the accuracy of the scanner depends on the environment in which it will be used. If the laser scanner is situated at a large z distance from the workspace, then the effects of error in the measurement of e will be reduced. This is the case in the CIRSSE testbed, since the distance between the scanning mirrors is 5 mm and the z distance to the workspace

is typically 2000 mm. If the laser scanner is used in situations where z is small, then the distance between the scanning mirrors should be determined analytically. This can be done in conjunction with determining the laser's extrinsic parameters (see Section 2.3).

Calibrating the scanning mirrors is critical to proper operation of the laser scanner. Each mirror is rotated with a galvanometer, which transforms a control voltage into an angular rotation of the mirror. Zero volts is assumed to correspond to a mirror angle of zero degrees (e.g., the laser beam is assumed to coincide with the laser's z axis when both galvanometers have zero input). The θ_y mirror is calibrated by directing the beam onto a plane at a fixed z distance from the laser with respect to the laser's origin (L). With θ_x fixed at zero, the θ_y mirror is rotated with a fixed voltage and the amount of y displacement on the plane is recorded. Using this information and (2.1) it is possible to determine θ_y. Assuming the relationship between voltage and mirror rotation is linear, mirror rotation is determined by dividing θ_y by the voltage applied to the mirror. To confirm the linearity of the galvanometers, the y mirror should be displaced to several different positions and the relationship should be verified to not change within measurement error. The θ_x mirror is calibrated in the same manner except that θ_y is fixed at zero (so there will be no pincushion distortion) and (2.2) is used to determine θ_x.

2.3 Calibration of the Extrinsic Parameters using an LSE Method

The extrinsic parameters of a laser scanner can be obtained using an LSE method as follows. Direct the laser at a set of 3-D points and record the scanning mirror angles at each point. These points and associated mirror angles can then be used to solve an overdetermined system of linear equations to obtain the laser's extrinsic parameters. The method described in this section is analogous to an LSE

approach proposed by Roger Tsai[12] for the calibration of cameras, except that the terms of the linear equations are different for a laser.

The extrinsic parameters of the laser scanner consist of the rotation and translation of the laser coordinate frame with respect to some other fixed coordinate frame. This rotation and translation should ultimately be represented as a 4×4 homogeneous transform of the type defined by Craig[13] of the form

$$ {}^{a}_{b}T = \begin{bmatrix} r_1 & r_2 & r_3 & t_x \\ r_4 & r_5 & r_6 & t_y \\ r_7 & r_8 & r_9 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.3} $$

The homogeneous transform ${}^{a}_{b}T$ is primarily composed of a 3×3 rotation matrix and a 3×1 translation vector which define the orientation and position of frame b with respect to frame a. To calibrate the laser, it is necessary to find ${}^{l}_{w}T$, which is the transformation from the laser coordinate frame to a desired world coordinate frame. What is required is a mathematical relationship that will determine these parameters given a set of points defined in the world coordinate frame and a set of corresponding scanning mirror angles. A point $\vec{P}_w$, defined in the world coordinate frame, is transformed to the laser coordinate frame using (2.4) to produce $\vec{P}_l$ ($\vec{P}_l$ and $\vec{P}_w$ are 3×1 vectors):

$$ \vec{P}_l = \begin{bmatrix} x_l \\ y_l \\ z_l \end{bmatrix} = {}^{l}_{w}T\vec{P}_w = {}^{l}_{w}T \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} = \begin{bmatrix} r_1x_w + r_2y_w + r_3z_w + t_x \\ r_4x_w + r_5y_w + r_6z_w + t_y \\ r_7x_w + r_8y_w + r_9z_w + t_z \end{bmatrix} \tag{2.4} $$

Also, x_l and y_l can be expressed in terms of the scanning mirrors using (2.1) and (2.2) with (2.4) as shown below:

$$ (z_l \sec\theta_y + e)\tan\theta_x = r_1x_w + r_2y_w + r_3z_w + t_x \tag{2.5} $$

$$ z_l \tan\theta_y = r_4x_w + r_5y_w + r_6z_w + t_y \tag{2.6} $$

Substituting the expression for z_w in (2.4) into (2.5) and (2.6) results in:

$$\begin{aligned}((r_7x_w + r_8y_w + r_9z_w + t_z)\sec\theta_y + e)\tan\theta_x &= \\ r_1x_w + r_2y_w + r_3z_w + t_x & \qquad (2.7)\\ (r_7x_w + r_8y_w + r_9z_w + t_z)\tan\theta_y &= \\ r_4x_w + r_5y_w + r_6z_w + t_y & \qquad (2.8)\end{aligned}$$

Simplifying (2.7) and (2.8) and then dividing through by t_z ($t_z \neq 0$):

$$\begin{aligned}\frac{\tan\theta_x}{\cos\theta_y}x_w\frac{r_7}{t_z} + \frac{\tan\theta_x}{\cos\theta_y}y_w\frac{r_8}{t_z} + & \\ \frac{\tan\theta_x}{\cos\theta_y}z_w\frac{r_9}{t_z} + \frac{\tan\theta_x}{\cos\theta_y} + \tan\theta_x\frac{e}{t_z} &= x_w\frac{r_1}{t_z} + y_w\frac{r_2}{t_z} + z_w\frac{r_3}{t_z} + \frac{t_x}{t_z} \qquad (2.9)\\ x_w\tan\theta_y\frac{r_7}{t_z} + y_w\tan\theta_y\frac{r_8}{t_z} + & \\ z_w\tan\theta_y\frac{r_9}{t_z} + \tan\theta_y &= x_w\frac{r1}{t_z} + y_w\frac{r_2}{t_z} + z_w\frac{r_3}{t_z} + \frac{t_y}{t_z} \qquad (2.10)\end{aligned}$$

Equations (2.9) and (2.10) can be expressed in the form $Ax = b$, where A is $2n \times 12$, b is $2n \times 1$, n is the number of data points collected, and x is a 12×1 vector of unknowns. Note that this system of equations not only determines the laser's extrinsic parameters, but also the distance between the scanning mirrors (e). The final form (for $t_z \neq 0$) is presented below:

$$\begin{bmatrix} \tan\theta_{x_1} & -x_{w_1} & -y_{w_1} & -z_{w_1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -x_{w_1} & -y_{w_1} & -z_{w_1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \tan\theta_{x_n} & -x_{w_n} & -y_{w_n} & -z_{w_n} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -x_{w_n} & -y_{w_n} & -z_{w_n} \end{bmatrix}$$

$$\begin{bmatrix} \frac{\tan\theta_{x_1}}{\cos\theta_{y_1}} x_{w_1} & \frac{\tan\theta_{x_1}}{\cos\theta_{y_1}} y_{w_1} & \frac{\tan\theta_{x_1}}{\cos\theta_{y_1}} z_{w_1} & -1 & 0 \\ x_{w_1}\tan\theta_{y_1} & y_{w_1}\tan\theta_{y_1} & z_{w_1}\tan\theta_{y_1} & 0 & -1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\tan\theta_{x_n}}{\cos\theta_{y_n}} x_{w_n} & \frac{\tan\theta_{x_n}}{\cos\theta_{y_n}} y_{w_n} & \frac{\tan\theta_{x_n}}{\cos\theta_{y_n}} z_{w_n} & -1 & 0 \\ x_{w_n}\tan\theta_{y_n} & y_{w_n}\tan\theta_{y_n} & z_{w_n}\tan\theta_{y_n} & 0 & -1 \end{bmatrix} - \begin{bmatrix} e \\ r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \\ r_6 \\ r_7 \\ r_8 \\ r_9 \\ t_x \\ t_y \end{bmatrix} \frac{1}{t_z} = \begin{bmatrix} \frac{\tan\theta_{y_1}}{\cos\theta_{x_1}} \\ \tan\theta_{y_1} \\ \vdots \\ \frac{\tan\theta_{y_n}}{\cos\theta_{x_n}} \\ \tan\theta_{y_n} \end{bmatrix} \tag{2.11}$$

After solving (2.11) using singular value decomposition [14], $|t_z|$ can be found from

$$|t_z| = \frac{1}{\sqrt{\left(\frac{r_7}{t_z}\right)^2 + \left(\frac{r_8}{t_z}\right)^2 + \left(\frac{r_9}{t_z}\right)^2}} \tag{2.12}$$

Once $|t_z|$ is determined, it is a straightforward process to obtain e, r_1, through r_9, t_x and t_y. With several hundred data points, this method produces reliable results provided that the data points are measured accurately (i.e., measurement error is ≤ 1 mm). The major flaw in this calibration method is that it treats the twelve parameters as being independent (which is obviously incorrect) and thereby fails to meet the constraints inherent in the rotation matrix.

2.4 Direct Geometric Method for Calibrating a Laser Scanner

Because the LSE calibration method generates a rotation matrix that is unconstrained, an alternative calibration method was developed that will produce an orthonormal rotation matrix. The method presented in this section treats laser

scanner calibration as a geometrical problem by taking advantage of the fact that the laser scanner emits a beam of light that can be measured with respect to a reference point. This method determines the laser scanner's orientation by directly determining its Euler angles. These Euler angles can then be used to produce a rotation matrix that is guaranteed to be orthonormal unlike the rotation matrix obtained using the LSE method. It should be emphasized that the objective of this calibration method is identical to that of the LSE method: to determine the homogeneous transform from the laser coordinate frame to some world coordinate frame.

The laser coordinate frame (L) is located at the center of the θ_y mirror and is oriented as in Figure 2.1. The pose of the world coordinate frame (W) is arbitrary. To assist in calibrating the laser, an intermediate coordinate frame F is defined whose origin is located directly below the center of the laser scanner's aperture. The frame F is located by suspending a plumb line from the laser scanner to some fixed plane (the floor is used in the CIRSSE system). This plumb line constitutes the z axis of the F coordinate frame. Note that while F's origin is on the floor, F's x and y axes are not necessarily coplanar with the floor. The location of F is selected in this manner to simplify the measurement of the laser scanner's orientation. The calibration procedure involves measuring twelve parameters, which are depicted in Figures 2.3 and 2.4 and defined as:

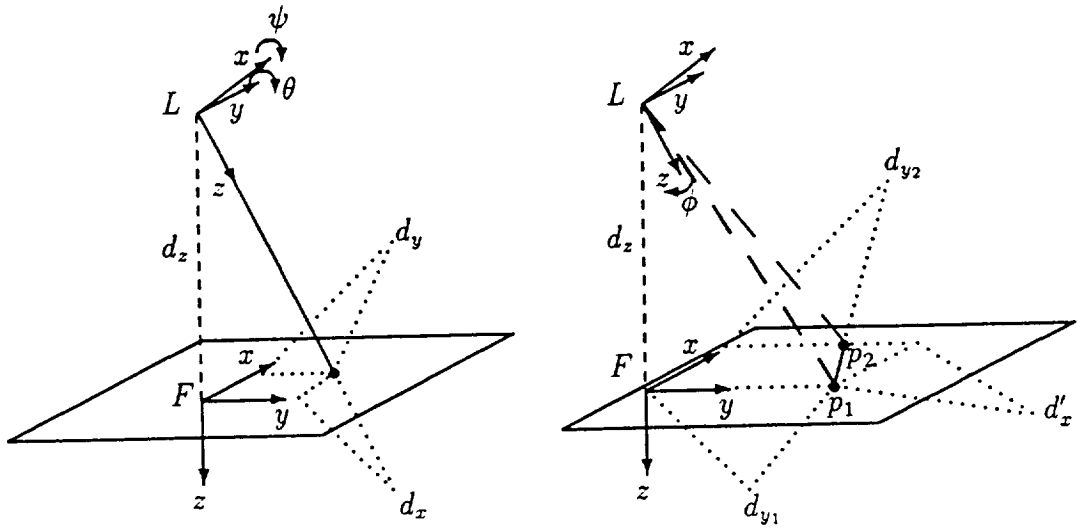


Figure 2.3: Determination of a laser scanner's euler angles. (Note: F is not necessarily coplanar with the floor).

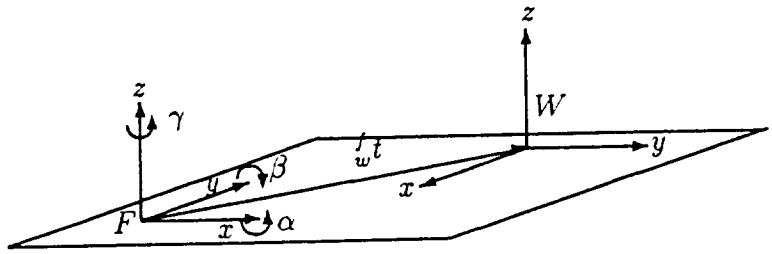


Figure 2.4: Transformation parameters from frame F to frame W. (Note: F is not necessarily coplanar with the floor).

d_z The distance from F to L along F's z axis.

d_x The x coordinate of the point where the undeflected laser beam intersects the floor, measured with respect to the origin of F.

d_y The y coordinate of the point where the undeflected laser beam intersects the floor, measured with respect to the origin of F.

d_{y_1} The y coordinate of the point p_1 on the floor, measured with respect to the origin of F.

d_{y_2} The y coordinate of the point p_2 on the floor, measured with respect to the origin of F.

d'_x The length of the projection of the line segment joining p_1 and p_2 onto the vector formed by projecting the x axis of frame F onto the plane of the floor.

ρ The rotation about F's x axis from F's xy plane to the floor.

δ The rotation about F's y axis from F's xy plane to the floor.

${}^f_w t$ The translation vector from F to W.

α The pitch angle about F's x axis from frame F to frame W

β The yaw angle about F's y axis from frame F to frame W

γ The roll angle about F's z axis from frame F to frame W

The above twelve parameters provide all the information required to determine the transformation from the laser coordinate frame to the world coordinate frame: ${}^l_w T$. The calculations are broken into two steps:

1. Determining ${}^l_f T$
2. Determining ${}^f_w T$

Once these transforms are known, it will be possible to determine ${}^l_w T$ from

$$ {}^l_w T = {}^l_f T \cdot {}^f_w T \tag{2.13} $$

${}^l_f R$ can be determined by deriving the orientation of frame F with respect to the laser in terms of the Euler angles pitch, yaw and roll (ψ,θ,ϕ) about the laser's x, y, and z axes, respectively. Figure 2.3 shows how the Euler angles are measured independently of one another. The laser's scanning mirrors are set to zero and the

coordinates of the point where the laser beam intersects the floor are measured (d_x, d_y). It is assumed that any offset of the scanning mirrors from zero is negligible compared to the magnitude of the Euler angles being measured. Since the mirrors are assumed to be in their undeflected state, the beam is coincident with the laser's z axis, and the point (d_x, d_y) is invariant to rotations about L's z axis. We will therefore assume that the laser coordinate frame is rotated about z by an amount ϕ (to be determined) so that the projection of the laser's y axis onto F's xy plane is coincident with F's y axis. Such an orientation implies that d_x is only a function of θ and d_y is only a function of ψ.

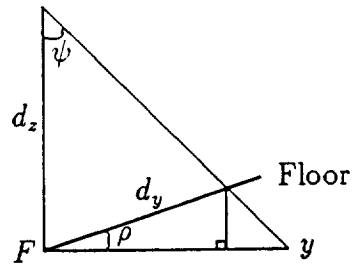


Figure 2.5: Relationship between the floor and F's xy plane

When calculating ψ and θ, it cannot be assumed that the floor correctly defines the xy plane of F. Indeed, since d_z was measured with a plumb bob, F's z axis is aligned with earth's local gravity vector, but the floor may not be orthogonal to this vector. By using a level it is possible to determine the angles ρ and δ about F's x and y axes, respectively, between the plane of the floor and the xy plane of F. The situation is depicted in Figure 2.5 for the Euler angle ψ. The formulas for determining ψ and θ can be derived directly from the figure and are presented below:

$$\psi = \tan^{-1}\left[\frac{d_y \cos\rho}{d_z - d_y \sin\rho}\right] \tag{2.14}$$

$$\theta = -\tan^{-1}\left[\frac{d_x \cos\delta}{d_z - d_x \sin\delta}\right] \tag{2.15}$$

At this point it is necessary to determine ϕ so that the laser's y axis can be aligned with F's y axis. The θ_x mirror is repeatedly rotated by an arbitrary amount while the θ_y mirror is set to zero, resulting in a line segment traced on the floor. The slope of this line determines ϕ and is found by measuring two arbitrary points (p_1 and p_2) on the line segment (this yields the parameters d_{y_1}, d_{y_2} and d'_x).

The primary issue with determining ϕ in this manner is ensuring the slope of the line segment to be a function only of ϕ and not of other variables, such as ψ, θ, or θ_y (which causes pincushion distortion). Because we have assumed $\theta_y = 0$, the slope of the line segment will not be altered by pincushion distortion. The Euler angle θ affects the laser's x and z components of the beam direction. A change in the z direction will not affect the slope of the line since the line ultimately will lie in F's xy plane. Further, distortion in x will tend to move the line segment by some constant value, leaving the slope unchanged. The Euler angle ψ affects the laser's y and z components of the beam direction. The distortion in z will not affect the measurements for the same reasons stated for θ. The distortion in y due to ψ consists of a constant translation of the line segment along F's y axis. The slope of the line is a function of the relative change in y from points p_1 to p_2, so the effects of ψ will not change the slope since p_1 and p_2 will be translated in y by the same amount. Hence, the slope of the line segment is a function only of ϕ.

Since the measurements used to determine ϕ were taken from the floor, they will have to be corrected for the effects of ρ and δ for the same reasons that the corrections were necessary for determining ψ and θ. The final equation for ϕ is presented below:

$$\phi = -tan^{-1}\left[\frac{(d_{y_2}\cos\rho) - (d_{y_1}\cos\rho)}{d'_x\cos\delta}\right] \tag{2.16}$$

At this point the Euler angles ψ, θ, ϕ have been determined, and it is now possible to create the rotation matrix ${}^{l}_{f}R$ from the laser to the F coordinate frame. Each Euler angle changes four parameters in the rotational matrix. The most straightforward approach is to determine separate rotational matrices for each Euler angle and then multiply them together to obtain ${}^{l}_{f}R$. The individual matrices for R(X,ψ), R(Y,θ) and R(Z,ϕ) are:

$$R(X,\psi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\psi & -\sin\psi \\ 0 & \sin\psi & \cos\psi \end{bmatrix} \tag{2.17}$$

$$R(Y,\theta) = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \tag{2.18}$$

$$R(Z,\phi) = \begin{bmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.19}$$

$${}^{l}_{f}R(\phi,\theta,\psi) = R(X,\psi)R(Y,\theta)R(Z,\phi) \tag{2.20}$$

As is apparent from (2.20), ${}^{l}_{f}R$ is obtained by combining the rotation matrices in the order roll, yaw, pitch. This ordering is essential for this calibration procedure. Applying the roll rotation first will align the projections of the laser's y axis to the F coordinate frame's y axis. This condition was assumed when the Euler angles ψ and θ were determined.

At this point, ${}^{l}_{f}R$ has been determined, but to obtain ${}^{l}_{f}T$ it is necessary to determine the translational component ${}^{l}_{f}t$ from the laser's origin to F with respect

to the laser. Given the configuration of the CIRSSE testbed and the available measuring equipment, it is difficult to directly measure this value with any degree of accuracy. However, it is not necessary to directly measure this value. The translation vector ${}^{f}_{l}t$ from F to the laser can be determined since F is located directly below the aperture of the laser, and d_z is known. The resulting value of ${}^{f}_{l}t$ is $[0, 0, d_z]^T$.

At this point ${}^{f}_{l}T$ can be determined by combining the inverse of ${}^{l}_{f}R$ (for a rotation matrix, $R^{-1} = R^T$) with the translation vector ${}^{f}_{l}t$. It is then possible to obtain ${}^{l}_{f}T$ by taking the inverse of ${}^{f}_{l}T$. This relationship is defined as

$$ {}^{l}_{f}T = \begin{bmatrix} & {}^{l}_{f}R^T & & {}^{f}_{l}t \\ & & & \\ 0 & 0 & 0 & 1 \end{bmatrix}^{-1} \tag{2.21} $$

With ${}^{l}_{f}T$ defined, the next step is to determine ${}^{f}_{w}T$. W should be chosen so that the Euler angles α, β, and γ can be easily measured. If W is chosen such that its z axis is plumb (as in the case with F), then the Euler angles from F to W can be readily determined. By applying (2.17) through (2.20), it is possible to determine ${}^{f}_{w}R$, and this matrix can be combined with ${}^{f}_{w}t$ (which is one of the twelve calibration parameters) to obtain ${}^{f}_{w}T$. At this point, ${}^{l}_{w}T$ can be determined by matrix multiplication:

$$ {}^{l}_{w}T = {}^{l}_{f}T \cdot {}^{f}_{w}T \tag{2.22} $$

This concludes the calibration of the laser's extrinsic parameters. It is now possible to transform points defined in the world coordinate frame into points defined in the laser coordinate frame. Additionally, points in the laser frame can be transformed into the world frame using the inverse of ${}^{l}_{w}T$.

Parameter	Value
d_x	5.0 mm
d_y	101.1 mm
d_z	2597.0 mm
ρ	0.0 radians
δ	0.0 radians
d_{y_1}	-10.0 mm
d_{y_2}	0.0 mm
dx'	946.0 mm
α	0.0 radians
β	π radians
γ	0.0 radians
${}^f_w t_x$	1409.9 mm
${}^f_w t_y$	887.65 mm
${}^f_w t_z$	0.0 mm
e	5.0 mm

Table 2.1: Typical calibration parameter values for direct geometric method

2.5 Appraisal of Calibration Method Performance

When presented with two alternatives for calibrating the laser's extrinsic parameters, the question arises as to which method is best. The answer depends on how and where the laser scanner will be used. The transform obtained using the direct geometrical method typically results in errors of less than 0.5%. This accuracy has consistently been obtained in the CIRSSE testbed where this method is currently implemented in software. Typical values for the calibration parameters are presented in table 2.1 This level of accuracy is sufficient for many visual sensing tasks. Further, the geometric approach achieves its results in a simple systematic manner. Hence, calibrating the laser using this method requires less effort than the LSE method.

Since the LSE method generates a solution based on a large set of data points, measurement errors among individual points should have less affect on the LSE solution. This is in contrast to the geometric method which uses a small number

of measurements to obtain its results, and, hence, these few points must be more accurately measured. The LSE method has been simulated in software and tested with simulated sets of data points containing differing degrees of error. The results of these simulations indicate that *if* it is possible to collect a large number of points with high accuracy, the LSE could produce more accurate results than the direct geometric method. At the time this research was conducted, the CIRSSE testbed had no means to collect a large number of highly accurate data points, but the LSE method has the potential to be highly effective when the testbed acquires the necessary data collection capability, and should be evaluated in the future.

3. Techniques Required for the Application of a Calibrated Laser and Camera to Visual Sensing

3.1 Overview

This chapter examines the development of methods to address the problems associated with applying a laser scanner and a camera to three dimensional visual sensing. As mentioned in section 1.2, these problems are:

1. Estimating three dimensional points using a calibrated camera and laser scanner.
2. Computer simulation models that provide a means to predict performance.
3. Methods to identify the laser spot in a cluttered camera image.
4. Methods of storing and representing three dimensional data.

Each problem is addressed individually. In some cases, more than one method is proposed for dealing with a specific problem. In such cases, the methods have been evaluated and compared to determine their merits and shortcomings. Additionally, experimental results are presented for these methods when appropriate or feasible. The information and conclusions presented here can be used when applying a laser scanner and a camera to three dimensional visual sensing. As will become readily apparent, application of these techniques to visual sensing tasks will depend heavily on the nature of the application.

3.2 Point Estimation with a Calibrated Laser and Camera

Once the laser scanner is calibrated to a world coordinate frame it is possible to use it in concert with a calibrated camera to perform three dimensional sensing.

Three dimensional sensing with a camera and a laser is different from strictly passive methods such as dual cameras. In a dual camera system, features identified in one camera image are correlated to similar features in the other camera image. The pixel coordinates of these features are then used to determine the three dimensional point of the object corresponding to the feature in the images.

Three dimensional sensing with a laser and camera, however, employs a slightly different approach. The laser directs its beam into the field of view of the camera and the camera image is scanned for the reflection of the laser beam off an object in the image. The pixel coordinates of the laser spot are correlated to the mirror angles of the laser scanner to obtain the three dimensional point of the object in the workspace.

Methods for estimating three dimensional points using dual cameras have been developed by Repko, Sood, and Kelly [15] and Noseworthy[16]. One method solves an overdetermined set of equations to obtain a lease squared estimate of the three dimensional point, while a second method calculates two three dimensional rays projecting from the image planes of the cameras, and estimates the coordinates of the corresponding point by determining the midpoint of the common normal of the rays.

These two methods for point estimation can be readily adapted for use with a calibrated camera and laser. In order to do this, the mathematical relationships contributed by one camera are substituted with mathematical relationships for the laser expressed in the same form as the camera. Hence, in order to prove that these point estimation methods will work with a laser and a camera, it is only necessary to show that the laser's mathematical model can be expressed in such a way as to be compatible with each method.

3.2.1 Point estimation using least squared error

The overdetermined system of linear equations approach for dual cameras[16] can readily be adapted to a camera and a laser scanner. What is needed is a relationship for the laser between the scanning mirror angles and the coordinates of a three dimensional point ($\vec{P}_w$) defined in the world frame. Such a relationship was derived in Section 2.3 as part of the LSE method for calibrating the laser (see equations (2.7) and (2.8)). Ultimately, the equations contributed by the laser will be included in a system of linear equations of the form $Ax = b$ where x is a three row by one column vector representing the x, y, and z coordinates of the estimated three dimensional point.

In the context of laser calibration, the unknown variables were the distance between the scanning mirrors (e) and the rotation and translation components of the transformation from the laser coordinate frame to the world coordinate frame. Since the laser is assumed to be calibrated at this point, all these values are known. Additionally, the angles of the scanning mirrors are known. What is not known are the x, y, and z coordinates of the world point. The terms in (2.7) and (2.8) can be regrouped into a form that is more suitable for point estimation:

$$\left(\frac{\tan\theta_x}{\cos\theta_y}r_7 - r_1\right)x_w + \left(\frac{\tan\theta_x}{\cos\theta_y}r_8 - r_2\right)y_w + \left(\frac{\tan\theta_x}{\cos\theta_y}r_9 - r_3\right)z_w = t_x - t_z\frac{\tan\theta_x}{\cos\theta_y} - e\tan\theta_x \quad (3.1)$$

$$(r_7\tan\theta_y - r_4)\,x_w + (r_8\tan\theta_y - r_5)\,y_w + (r_9\tan\theta_y - r_6)\,z_w = t_y - t_z\tan\theta_y \quad (3.2)$$

Therefore, the laser contributes two equations with the same three unknowns as the camera equations (assuming that the laser and camera are calibrated to the same world coordinate frame). These two equations can be combine with the two

equations contributed by the camera[16]. For sake of brevity, the full derivation of the camera's equations will not be shown here, but the final result is given by equations (3.3) and (3.4).

$$\begin{aligned} x_w(r_7x_u - fr_1) + y_w(r_8x_u - fr_3) + z_w(r_9x_u - fr_3) &= \\ ft_x - t_zx_u & \qquad (3.3) \\ x_w(r_7y_u - fr_4) + y_w(r_8y_u - fr_5) + z_w(r_9y_u - fr_6) &= \\ ft_y - t_zy_u & \qquad (3.4) \end{aligned}$$

Where x_u and y_u are camera pixel coordinates, f is the focal length of the camera, and r_1 - r_9, t_x, t_y and t_z are elements of the homogeneous transformation ${}^c_wT$ from the camera's coordinate space to the world coordinate space. Equations (3.3), (3.4), (3.1), and (3.2) can be expressed as a system of linear equations of the form $Ax = b$:

$$\begin{bmatrix} (r_7x_u - fr_1) & (r_8x_u - fr_3) & (r_9x_u - fr_3) \\ (r_7y_u - rr_4) & (r_8y_u - fr_5) & (r_9y_u - fr_6) \\ \left(\frac{\tan\theta_x}{\cos\theta_y}r_7 - r_1\right) & \left(\frac{\tan\theta_x}{\cos\theta_y}r_8 - r_2\right) & \left(\frac{\tan\theta_x}{\cos\theta_y}r_9 - r_3\right) \\ (r_7\tan\theta_y - r_4) & (r_8\tan\theta_y - r_5) & (r_9\tan\theta_y - r_6) \end{bmatrix} \cdot \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} = \begin{bmatrix} (ft_x - t_zx_u) \\ (ft_y - t_zy_u) \\ \left(t_x - t_z\frac{\tan\theta_x}{\cos\theta_y} - e\tan\theta_x\right) \\ (t_y - t_z\tan\theta_y) \end{bmatrix} \qquad (3.5)$$

The system of equations presented in (3.5) can be solved using singular value decomposition[14]. Care must be taken, however, in interpreting the results of the

LSE solution[16]. The LSE approach solves for x by minimizing the expression: $(b - Ax)^T (b - Ax)$ the solution to which is $x = \left(A^T A\right)^{-1} \cdot A^T b$ and the quantity $\left(A^T A\right)^{-1}$ is the pseudo-inverse of A. Since this is "implicit" estimation, minimizing $(b - Ax)^T (b - Ax)$ does not mean that the error between the actual point and the estimated point has been minimized. The error $(b - Ax)^T (b - Ax)$ is minimized, and by so doing, it is assumed that the parameter of interest (x) is optimized in the process, which is generally true.

3.2.2 Point estimation using midpoint of common normal

Another method for estimating three dimensional points was developed by Noseworthy[16]. This method calculates a ray projecting from the camera's image plane into the three dimensional environment. A brief summary of this work is presented below followed by a description of how this method can be adapted for use with a laser.

A three dimensional point is calculated by determining the midpoint to the common normal of the rays calculated from two different cameras. The ray for each camera is expressed as a linear parametric equation of the form:

$$\vec{r}_c = s_c \vec{d}_c + \vec{O}_c \tag{3.6}$$

Where $\vec{O}_c$ is the 3×1 vector describing the location of the origin of the camera with respect to the world coordinate system, and $\vec{d}_c$ is a 3×1 vector describing the direction of the ray projected from the camera's image plane with respect to the world coordinate system. These two terms can be further expressed as:

$$\vec{d}_c = {}^w_c R \begin{bmatrix} x_u \\ y_u \\ f_c \end{bmatrix} s_c \tag{3.7}$$

$$\vec{O}_c = {}^w_c t \tag{3.8}$$

${}^w_c R$ and ${}^w_c t$ can be obtained from the inverse of the homogeneous transform ${}^c_w T$, and x_u and y_u are camera pixel coordinates of a pinhole camera model. Once these parametric equations have been determined it is possible to calculate the unit direction, in world coordinates, of the common normal to the rays from two cameras as:

$$\hat{n} = \frac{\vec{d}_{c_1} \times \vec{d}_{c_2}}{| \vec{d}_{c_1} \times \vec{d}_{c_2} |} \qquad | \vec{d}_{c_1} \times \vec{d}_{c_2} | \neq 0 \tag{3.9}$$

The shortest distance between $\vec{r}_{c_1}$ and $\vec{r}_{c_2}$, l, can be determined by projecting $\vec{O}_{c_1} - \vec{O}_{c_2}$ in the $\hat{n}$ direction. Noseworthy points out that $\vec{r}_{c_1}$ and $\vec{r}_{c_2}$ are assumed to be skew (i.e. $| \vec{d}_{c_1} \times \vec{d}_{c_2} | \neq 0$). l is determined using the following expression:

$$l = \frac{\left(\vec{O}_{c_1} - \vec{O}_{c_2}\right) \cdot \left(\vec{d}_{c_1} \times \vec{d}_{c_2}\right]}{| \vec{d}_{c_1} \times \vec{d}_{c_2} |} \tag{3.10}$$

The next step is to solve for s_1 and s_2. This can be done using the following relationship:

$$\left(s_1 \vec{d}_{c_1} + \vec{O}_{c_1}\right) - \left(s_2 \vec{d}_{c_2} + \vec{O}_{c_2}\right) = l\hat{n} \tag{3.11}$$

Finally, a 3×1 vector, $\vec{m}$, representing the coordinates of the midpoint to the common normal of $\vec{r}_{c_1}$ and $\vec{r}_{c_2}$ is determined by:

$$\vec{m} = r_{c_1} \mid_{s_1} -\frac{1}{2} l\hat{n} = r_{c_2} \mid_{s_2} +\frac{1}{2} l\hat{n} \tag{3.12}$$

Where $r_{c_1} \mid_{s_1}$ and $r_{c_2} \mid_{s_2}$ represent the parametric equations for the rays projected from cameras one and two evaluated at s_1 and s_2 respectively.

The midpoint to the common normal method can also be used with a calibrated camera and laser scanner. To do this, it is necessary to derive a parametric equation for the 3-D ray of the laser beam. This equation is expressed as

$$\vec{r_l} = s_l \vec{d_l} + \vec{O_l} \tag{3.13}$$

where $\vec{O_l}$ is the origin of the laser beam in terms of the world coordinate frame, $\vec{d_l}$ is the direction vector of the laser ray and s_l is a parameter.

Equation (3.13) is of the same form as (3.6). To use the midpoint to the common normal method with a calibrated laser, it is necessary to derive expressions for $\vec{O_l}$, $\vec{d_l}$ and s_l. Once these values are determined, the mathematical relationships for the midpoint of the common normal for two cameras will also work for a calibrated camera and laser.

$\vec{O_l}$ can be determined based on the values of the intrinsic and extrinsic parameters of the laser scanner. Specifically, the required parameters are the distance between the scanning mirrors, and the transformation ${}^{l}_{w}T$ from the laser coordinate frame to the world coordinate frame. The laser calibration assumes the laser's origin to be at the center of the θ_y mirror. The coordinates of the laser's origin with respect to the world coordinate frame can be derived by extracting the translational component of the inverse of ${}^{l}_{w}T$. This translation vector is defined as ${}^{w}_{l}t$.

The origin of the laser has one additional component. Recalling the arrangement of the laser scanner described in Section 2.1, the θ_x mirror deflects the laser

beam along the rotational axis of the θ_y mirror. Since the origin of the laser scanner is defined to be on the θ_y mirror, the origin of the laser ray is translated along the x axis of the laser scanner (this is the rotational axis of the θ_y mirror) by the rotation of the θ_x mirror. The translation of the laser's origin as a function of θ_x with respect to the laser's coordinate frame can be expressed as a 3×1 vector:

$$\vec{t}_{\theta_x} = \begin{bmatrix} e \tan \theta_x \\ 0 \\ 0 \end{bmatrix} \tag{3.14}$$

The term $e \tan \theta_x$ from (2.2) defines the x coordinate of the laser beam with respect to the laser's coordinate frame given a set of mirror angles and a specific z coordinate. The vector $\vec{t}_{\theta_x}$ is defined in terms of the laser coordinate frame; hence it must be transformed into a translation with respect to the world coordinate frame. This is accomplished by multiplying $\vec{t}_{\theta_x}$ by the rotation matrix ${}^w_l R$ contained in the inverse of the homogeneous transform ${}^l_w T$. This yields a new translation vector defined in the world coordinate frame:

$$\vec{t_{w_{\theta_x}}} = {}^w_l R \cdot \vec{t_{\theta_x}} \tag{3.15}$$

Therefore, the final value for $\vec{O}_l$ can be expressed as

$$\vec{O}_l = {}^w_l t + \vec{t}_{w_{\theta_x}} \tag{3.16}$$

The next step is to determine $\vec{d}_l$ from (2.2) and (2.1). Since the expression for $\vec{O}_l$ already compensates for the translation of the laser's origin due to rotation of the θ_x mirror, the $e \tan \theta_x$ term in (2.2) can be removed. Hence, the direction of the laser ray with respect to the laser coordinate frame can be expressed as a 3×1

vector of the form

$$\vec{d_l} = \begin{bmatrix} z \sec\theta_y \tan\theta_x \\ z \tan\theta_y \\ z \end{bmatrix} \tag{3.17}$$

The direction vector is currently defined with respect to the laser coordinate frame. In order to use it in determining the midpoint of the common normal, it will have to be transformed with respect to the world coordinate frame. This can be done simply by multiplying $\vec{d_l}$ by the rotation matrix ${}^{w}_{l}R$. Further, the z term in $\vec{d_l}$ can be factored out and used as the variable parameter s_l. Therefore, the parametric equation for the 3-D laser ray can be expressed as

$$\vec{r_l} = {}^{w}_{l}R\vec{d_l}z + \vec{O_l} \tag{3.18}$$

Equation (3.18) can be used in place of the parametric equation for the second camera to determine the midpoint of the common normal for a camera and a laser scanner. The remaining mathematical expressions for the midpoint of the common normal calculation remain valid.

3.2.3 Appraisal of point estimation methods

As mentioned previously, the LSE point estimation method generates a solution by attempting to optimize the value of x. Because this method attempts to determine an optimized solution, it can accommodate minor errors in the camera and laser calibration parameters. This property of the LSE method can be useful in that minor calibration errors will not necessarily result in poor point estimates. Indeed, tests results have indicated that the LSE method can provide accurate results even if the rotation matrix is not truly orthonormal.

The midpoint to the common normal method attempts to model the exact behavior of the laser and camera. Specifically, it projects rays from the laser and the camera into space based on the mirror angles of the laser scanner and the pixel coordinates on the camera's image plane. If all the calibration parameters for the laser and camera are perfectly accurate, the rays should intersect, but in reality, the rays do not intersect due to errors in calibration, and hence, the midpoint of the common normal to these two rays is used as the estimate of the three dimensional point.

The decision as to which method to use for point estimation depends on the nature of the application in which the method will be used. Since the LSE method is more robust in terms of accommodating calibration errors it may be useful in situations where the accuracy of calibration parameters, particularly the orthogonality of the rotation matrix, is questionable.

The midpoint to the common normal method does have one significant quality. Since this method is a direct representation of the geometry of the point estimation scenario, it may be possible to use this method as a means of predicting how calibration errors will affect performance. The comparative performance of the LSE and midpoint of the normal point estimation methods is a topic that is addressed in the simulation models presented in the next section.

3.3 Computer Simulation Models of a Laser Scanner

Ideally, the laser scanner should operate "perfectly" assuming it is calibrated to a degree where no error exists or is so small as to be unobservable. Unfortunately, such a situation is rarely the case. It is more likely that the laser scanner's calibration will exhibit some inaccuracies and therefore it is important to understand how such errors will degrade performance. Additionally, if the laser scanner is ever

mounted on a platform that is prone to vibration or other extraneous motion, the laser's calibration (especially the extrinsic parameters) is sure to change. In short, it would be advantageous to be able to simulate the effect of calibration errors as a means of predicting laser scanner performance and possibly providing a means of compensating for inaccuracy.

Two computer simulations are presented. The first model simulates the inaccuracy associated with errors in the laser scanner's calibration. A second simulation examines the effect of laser and camera point estimation given that either the camera or the laser scanner are perturbed.

3.3.1 Simulation of laser scanner calibration

In order to develop a simulation for laser scanner calibration, it is necessary to answer three basic questions. First, what calibration method should be simulated? Second, how is *accuracy* or *error* measured? Third, how will the performance of the laser scanner be evaluated? Once these questions are answered, generating a simulation is straightforward.

Chapter 2 presented two methods for calibrating a laser scanner's extrinsic parameters. Each methods has its merits and limitations as previously discussed, but it is important to identify which method (or both) should be simulated. Solutions generated by the LSE solution are highly dependent on the number of data points taken and the error associated with these points. Therefore it is difficult to assess how a specific error in a single data point (or a small set of points) will affect the calibration parameters generated by the LSE method. Hence, it is not well suited for simulation.

The direct geometric calibration method has several qualities that make it amenable to simulation modeling. First, the extrinsic parameters generated by this

method are directly dependent on a finite set of input parameters. Second, the effect of each input parameter on the extrinsic parameters is well defined using the formulas presented in section 2.4. It is therefore possible to perturb an input parameter and obtain a repeatable change in the extrinsic parameters. Third, the direct geometric method is currently used to calibrate the laser scanner, so simulating it would directly benefit the ongoing vision research in the CIRSSE testbed. Fourth, the direct geometric method models the inherent geometry of the laser scanner. For these reasons, the simulation uses the direct geometric calibration model.

When predicting the accuracy of the laser calibration it is important to define exactly what is meant by accuracy. Accuracy is defined to be how well the laser scanner is capable of directing the laser beam at a specific three dimensional point defined in the world coordinate system (P_w). To measure this accuracy, it will be necessary to create a gauge of how far away the laser beam will be from the desired point. The error measure used in this simulation is the distance of the normal from the laser beam to P_w. To calculate this, the laser beam is expressed as a parametric equation of the form shown in equation (3.13). A point Q (see figure 3.1) is defined at the intersection of the laser beam with the line normal to the laser beam and passing through P_w. The vector $\vec{P_wQ}$ can be determined as:

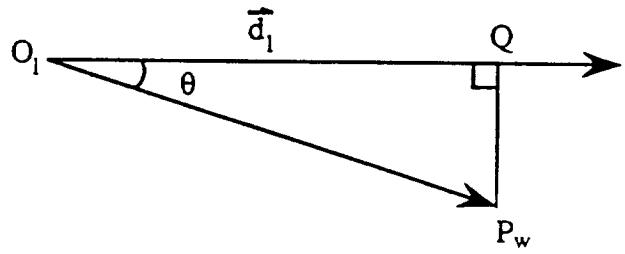


Figure 3.1: Calculation of error vector from target point to laser beam

$$\vec{P_w Q} = -\vec{O_l P_w} + \frac{\vec{d_l}}{|\vec{d_l}|} |\vec{O_l P_w}| \cos\theta \tag{3.19}$$

It is generally known that:

$$\vec{O_l P_w} \cdot \frac{\vec{d_l}}{|\vec{d_l}|} = |\vec{O_l P_w}| \, |\frac{\vec{d_l}}{|\vec{d_l}|}| \cos\theta \tag{3.20}$$

however, $|\frac{\vec{d_l}}{|\vec{d_l}|}| = 1$ therefore, by substitution:

$$\vec{P_w Q} = -\vec{O_l P_w} + \frac{\vec{d_l}}{|\vec{d_l}|} \left(\vec{O_l P_w} \cdot \frac{\vec{d_l}}{|\vec{d_l}|} \right) \tag{3.21}$$

For purposes of analysis, the magnitude of $\vec{P_w Q}$ can be examined to determine the distance of the error, or the individual x, y, z components can be examined if directional information is required.

The final step in generating the simulation is to determine a method by which the laser scanner calibration will be evaluated. This analysis can be performed in the following manner:

1. Define a matrix of test points defined in world coordinates. This matrix should define some plane in the workspace.

2. Using a given set of input parameters (d_z, d_x, d_y etc.) as defined in section 2.4 calculate ${}^{l}_{w}T$ using the direct geometric calibration method.

3. Using ${}^{l}_{w}T$ and equation (2.4) calculate the mirrors angles required to direct the laser beam to each of the points in the matrix.

4. Perturb the input parameters and calculate ${}^{l}_{w}T'$ using the direct geometric method.

5. Using ${}^{l}_{w}T'$ and the pairs of mirror angles previously calculated, determine r'_l (the parametric equation of the laser beam) for each pair of mirror angles.

6. Calculate $\vec{P_wQ}$ for each P_w in the matrix and its respective r'_l. The set of $\vec{P_wQ}$ values denotes the accuracy of the laser scanner for the points in the plane.

7. Repeat the steps stated above for different planes to obtain the accuracy of the laser scanner throughout the workspace.

The simulation constructed for this research generates a matrix of points that lie in a plane that is parallel to the XY plane of the world coordinate system at some fixed world Z. In effect, the simulation is designed to take "slices" of CIRSSE's testbed from its floor to its ceiling. Although the planes are defined as parallel to the worlds XY plane, there is no reason why this can't be modified to accommodate some other plane as long as a transformation exists from this plane to world coordinates. The laser scanner simulation is designed to generate output that is compatible with analysis programs such as Matlab, Delta Graph or PvWave. These file formats are supported because Matlab, Delta Graph and PvWave are most readily available to the author and to CIRSSE and otherwise popularly in use.

3.3.2 Results of laser calibration simulation

The laser calibration simulation was subjected to a variety of tests to evaluate the effect of perturbing the individual input parameters to the direct geometric method. Table 3.1 shows the values of the input parameters before the perturbations were applied. A total of fifteen tests were performed. Table 3.2 indicates for each test, which parameter was altered and by how much.

Each test was conducted at three different planes in the CIRSSE testbed. The dimensions of these planes are shown in table 3.3. Each plane had four hundred points distributed evenly as a 20×20 grid. The coordinates of these planes coincide with a section of the workspace in the CIRSSE testbed accessible to both PUMA

Parameter	d_x	d_y	d_z	ρ	δ
Value	45mm	-6mm	25910mm	0.0rad	0.0rad
Parameter	d_{y1}	d_{y2}	d'_x	α	β
Value	30mm	40mm	1314mm	-0.0rad	π
Parameter	γ	${}^f_w t_x$	${}^f_w t_y$	${}^f_w t_z$	e
Amount Changed	0.0rad	1413mm	33mm	0.0mm	5mm

Table 3.1: Laser calibration parameters at start of tests

Test	1	2	3	4	5
Parameter	d_x	d_y	d_z	ρ	δ
Amount Changed	10mm	3mm	-20mm	0.2rad	0.1rad
Test	6	7	8	9	10
Parameter	d_{y1}	d_{y2}	d'_x	α	β
Amount Changed	5mm	-10mm	-10mm	-0.1rad	0.3rad
Test	11	12	13	14	15
Parameter	γ	${}^f_w t_x$	${}^f_w t_y$	${}^f_w t_z$	e
Amount Changed	0.1rad	10mm	8mm	-10mm	5mm

Table 3.2: Tests performed using laser calibration simulation

robots. For each test at each plane, the x, y, z and magnitude of $\vec{P_w Q}$ was recorded. The tests generated approximately 150 different error graphs. The more interesting plots are included in this report while general observations will be made about the results of the other tests.

Plane	Min X	Min Y	Max X	Max Y	Z
1	1000mm	-750mm	2500mm	750mm	0mm
2	1200mm	-500mm	2200mm	500mm	300mm
3	1200mm	-500mm	2200mm	500mm	600mm

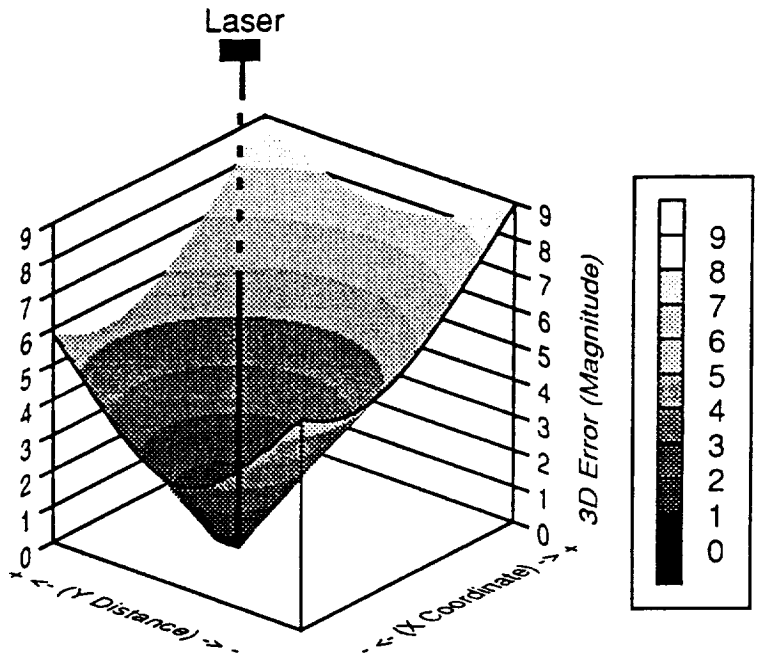
Table 3.3: Dimensions of point planes (in world coordinates)

The tests that deserve particular attention are those that perturbed the d_z, d_x and α parameters. The magnitude of the error induced by perturbing the d_z parameter is shown in figure 3.2. The projection of the laser's origin onto the XY plane of the testbed is indicated for reference. Note that the laser is *not* centered

over the workspace, especially in the X axis. This was intentionally done to better represent the actual conditions of the CIRSSE testbed where the laser is not physically centered over the workspace of the robots. The figure shows the error at z = 0mm. The error curves were the same at z = 300mm and z = 600mm, but the overall error was progressively smaller as z increased. The shape of the curve is not surprising when one considers that the d_z parameters affects the Euler angles ψ and θ as well as the translation vector ${}^{l}_{f}t$. The sum effect of perturbing all these parameters is the conical shape in the figure. The magnitude of the error increases symmetrically about the origin of the laser. This is revealed in figure 3.3 which shows the projection of the error regions of figure 3.2 onto the XY plane of the world coordinate system. The projection of the laser's origin is indicated as a white cross. The perturbation of d_z is a simple condition to recreate in the CIRSSE testbed. When the d_z was perturbed in the testbed, the same type of error behavior was observed.

A second test that produced interesting results was the perturbation of the d_x parameter. This parameter only affects the Euler angle θ. A three dimensional surface plot and a contour plot (error projected onto the world XY plane) are presented in figures 3.4 and 3.5. The error is relatively constant at 9.0mm, but while the value is constant along the y axis, it is slightly distorted in the x axis. This distortion is reasonable considering that θ is a rotation about laser's y axis.

The most dramatic results of calibration model were observed in the perturbation of the α parameter. The resulting plots were not only non-linear, but also were significantly different as z increased. The three dimensional surface and contour plots for these tests are shown in figures 3.6 - 3.11. As is apparent in figure 3.6, when z = 0mm the error reaches a minimum directly beneath the laser scanner and the error becomes progressively larger. Further, the error increases as the



X = [1000, 2500] Y=[-750, 750] Z = 0

(all units in millimeters)

Figure 3.2: Surface plot of the magnitude of error caused by perturbing d_z

distance along the world y axis from the point to the laser's origin increases, yet is relatively constant as the distance along the world x axis increases. The constant behavior along the x axis is reasonable since α is a rotation about the x axis of the F coordinate frame.

As previously mentioned, the error curves for the d_z and d_x parameters exhibited the same behavior regardless of the values of z. It is therefore reasonable to have assumed that this phenomenon would appear when perturbing the α parameter. In reality, however, the error curve at $z = 300$mm is different from the error curve at $z = 0$mm. Indeed, the error achieves a maximum value directly under the laser. Because the results were so unusual, an additional test run was conducted using a

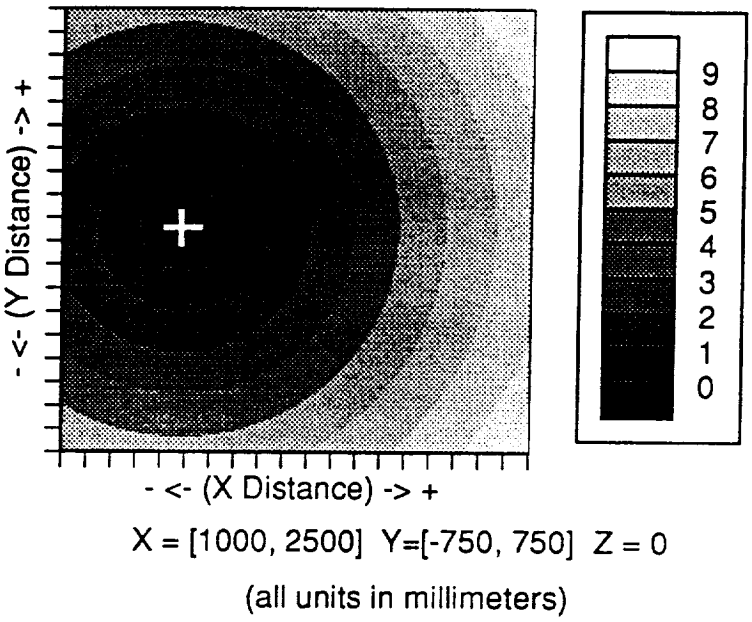


Figure 3.3: Contour plot of the magnitude of error caused by perturbing d_z

plane of test points at $z = 150$mm. The three dimensional surface and contour plots of the error for this test are shown in figures 3.8 and 3.9 respectively. These curves indicate that the error curve behavior is in a transition from the $z = 0$mm curve to the $z = 300$mm curve. This unusual behavior deserves more detailed examination.

It is necessary to define the 3×1 vector P_w which is a point in the test plane defined in world coordinates. In order to direct the laser to this point it is first transformed to the laser's coordinate system by applying the transform ${}^l_w T$:

$$P_l = {}^l_w T \cdot P_w \tag{3.22}$$

By applying equation 2.22 the expression becomes:

$$P_l = {}^l_f T \cdot {}^f_w T \cdot P_w \tag{3.23}$$

Expanding ${}^f_w T$ using the rotations $R(X, \alpha)R(Y, \beta)R(Z, \gamma)$ as described section

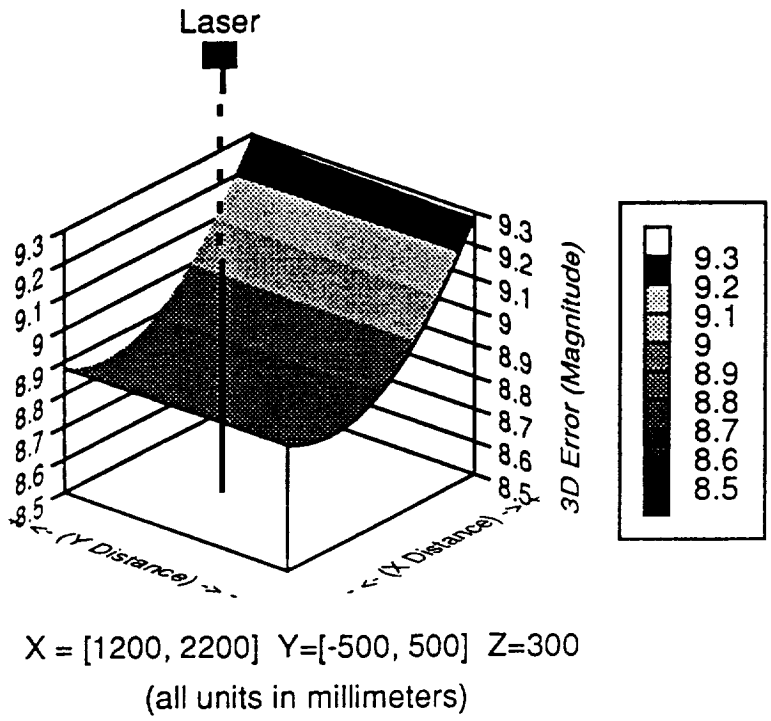


Figure 3.4: Surface plot of the magnitude of error caused by perturbing d_x

2.4 yields:

$$P_l = {}^{l}_{f}T \cdot \begin{bmatrix} c\gamma c\beta & -s\gamma c\beta & s\beta & {}^{f}_{w}t_x \\ c\gamma s\beta s\alpha + s\gamma c\alpha & -s\gamma s\beta s\alpha + c\gamma c\alpha & -c\beta s\alpha & {}^{f}_{w}t_y \\ -c\gamma s\beta c\alpha + s\gamma s\alpha & s\gamma s\beta c\alpha + c\gamma s\alpha & c\beta c\alpha & {}^{f}_{w}t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot P_w \qquad (3.24)$$

Where c is cos and s is sin. From table 3.1 $\gamma = 0.0$ and $\beta = \pi$. Substituting these values into the matrix and reducing terms results in:

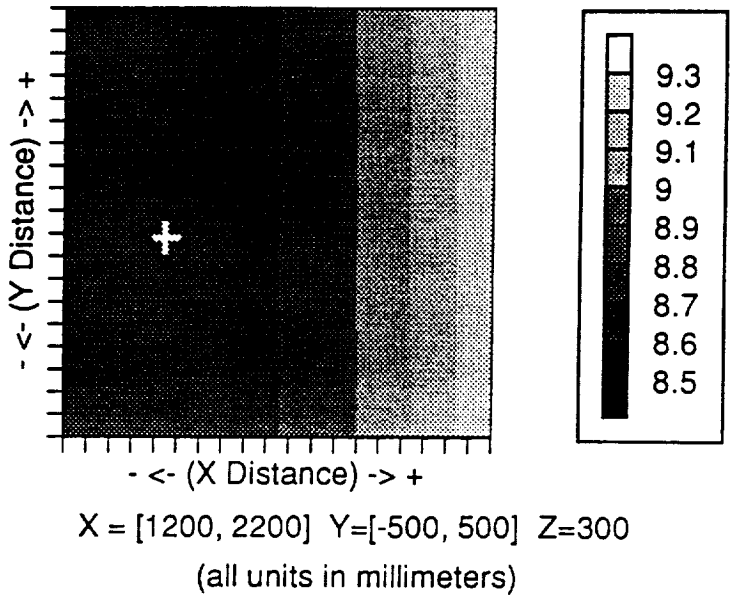


Figure 3.5: Contour plot of the magnitude error caused by perturbing d_x

$$P_l = {}^l_f T \cdot \begin{bmatrix} -1 & 0 & 0 & {}^f_w t_x \\ 0 & c\alpha & s\alpha & {}^f_w t_y \\ 0 & s\alpha & -c\alpha & {}^f_w t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot P_w \tag{3.25}$$

Since none of the parameters used to calculate ${}^l_f T$ are changed, the terms inside this transform are constants. Expanding ${}^l_f T$ gives:

$$P_l = \begin{bmatrix} r_1 & r_2 & r_3 & {}^l_f t_x \\ r_4 & r_5 & r_6 & {}^l_f t_y \\ r_7 & r_8 & r_9 & {}^l_f t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 0 & 0 & {}^f_w t_x \\ 0 & c\alpha & s\alpha & {}^f_w t_y \\ 0 & s\alpha & -c\alpha & {}^f_w t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot P_w \tag{3.26}$$

Simplifying the equation and solving for P_l yields:

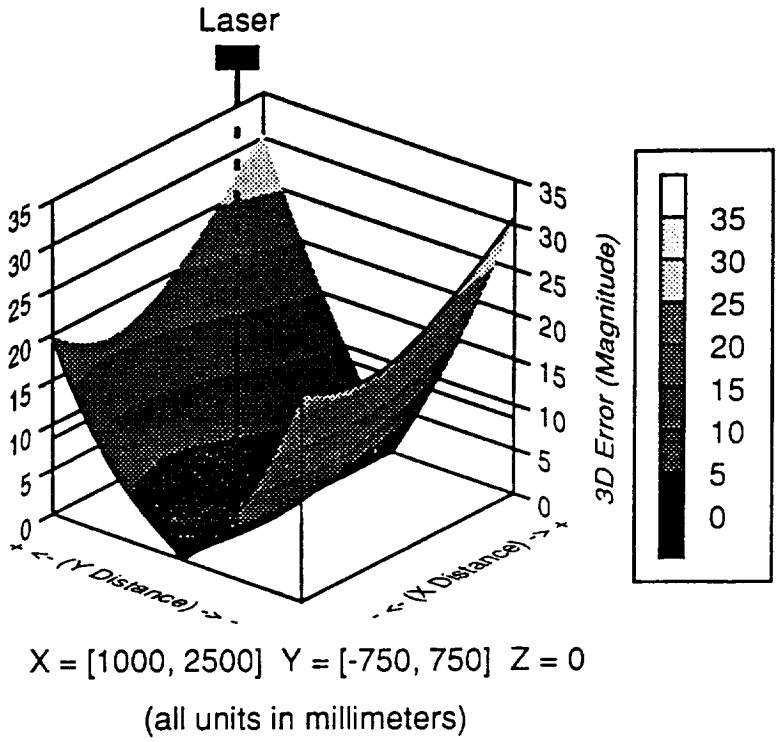


Figure 3.6: Surface plot of the magnitude error caused by perturbing the α parameter. z = 0mm

$$P_l = \begin{bmatrix} r_1(-P_{w_x}+{}^f_w t_x)+r_2(P_{w_y}c\alpha+P_{w_z}s\alpha+{}^f_w t_y)+r_3(P_{w_y}s\alpha-P_{w_z}c\alpha+{}^f_w t_z)+{}^i_f t_x \\ r_4(-P_{w_x}+{}^f_w t_x)+r_5(P_{w_y}c\alpha+P_{w_z}s\alpha+{}^f_w t_y)+r_6(P_{w_y}s\alpha-P_{w_z}c\alpha+{}^f_w t_z)+{}^i_f t_y \\ r_7(-P_{w_x}+{}^f_w t_x)+r_8(P_{w_y}c\alpha+P_{w_z}s\alpha+{}^f_w t_y)+r_9(P_{w_y}s\alpha-P_{w_z}c\alpha+{}^f_w t_z)+{}^i_f t_z \end{bmatrix} \quad (3.27)$$

The behavior observed in the surface plots is due to the $(P_{w_y}s\alpha - P_{w_z}c\alpha + {}^f_w t_z)$ terms in the expression. When Z = 0, the value of $P_{w_z}c\alpha$ is zero. Hence, the entire term increases as a function of P_{w_y}. When z is not zero, $P_{w_z}c\alpha$ is non-zero. As P_{w_y} increases, the term approaches zero and becomes increasingly negative. Since, the plot represents the absolute magnitude of the error, the term creates the appearance of approaching zero and then increasing in value. This entire process is clearly shown in the case of z = 150mm, but there is no evidence of it at z = 300mm. This is due

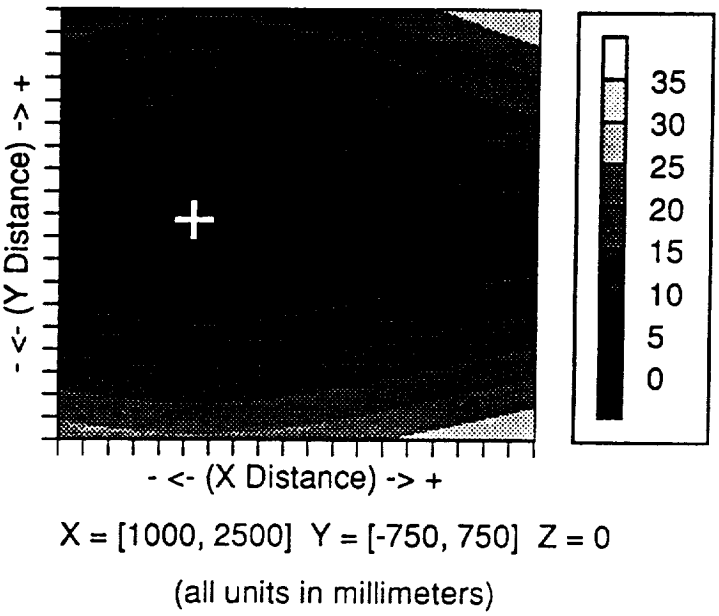


Figure 3.7: Contour plot of the magnitude error caused by perturbing the α parameter. z = 0mm

to the fact that the dimensions of the plane of test points is not large enough to show the entire process at z=300mm. If the dimensions of the plane are increased, the behavior exhibited at z=150 is evident at z = 300mm.

The error plots at z = 150mm also indicate that the two regions of small error occur near the projected origin of the laser on the world xy plane. This is attributable to the fact that P_{w_x} is small in this region. As P_{w_x} becomes large, the behavior of the $(P_{w_y} s\alpha - P_{w_z} c\alpha + \frac{f}{w} t_z)$ term becomes less significant.

The unusual behavior exhibited by the α parameter is due, in part, to the specific parameters used in the simulation. If different parameters were used, the α parameter might behavior in a more consistent manner. Since there is a potential for such unpredictable results, the utility of the laser calibration simulation is readily apparent.

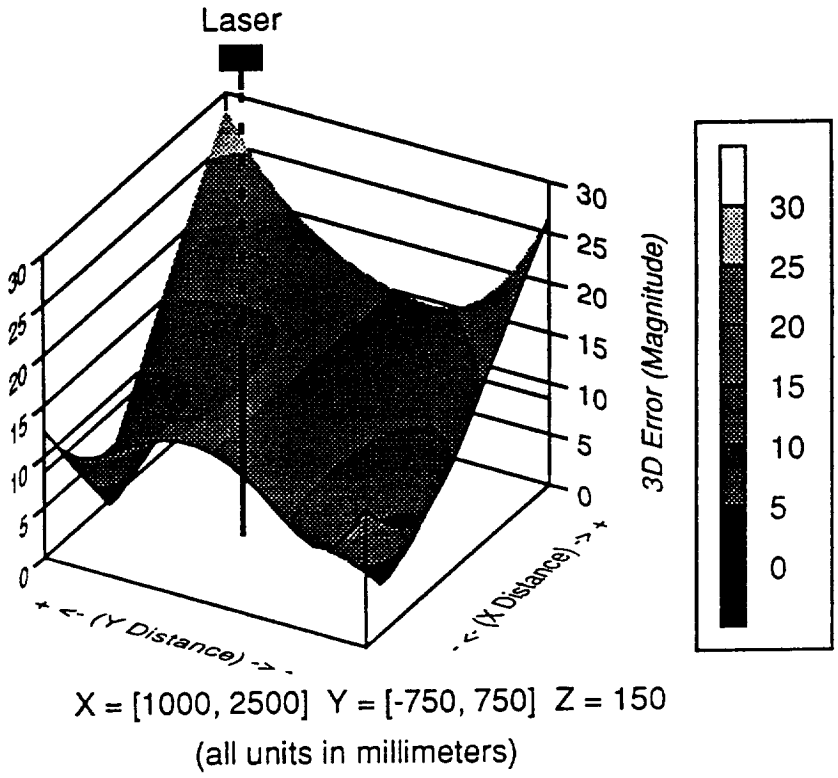


Figure 3.8: Surface plot of the magnitude error caused by perturbing the α parameter. z = 150mm

The results of the other tests yield some general observations about the laser's performance given the specific conditions used for the test. Since the test conditions are representative of conditions in the CIRSSE testbed, the observations also apply to it. These observations are:

- Perturbations in calibration parameters that affect the rotational components of the laser's pose exhibited more erratic error than perturbations that affect the translational components. This makes sense since a rotation affects two coordinate axes and does so as a nonlinear function (i.e. sin and cos), and the amount of error due to a rotational shift will depend on how far away from the origin a point is situated. A translational shift, however is a constant

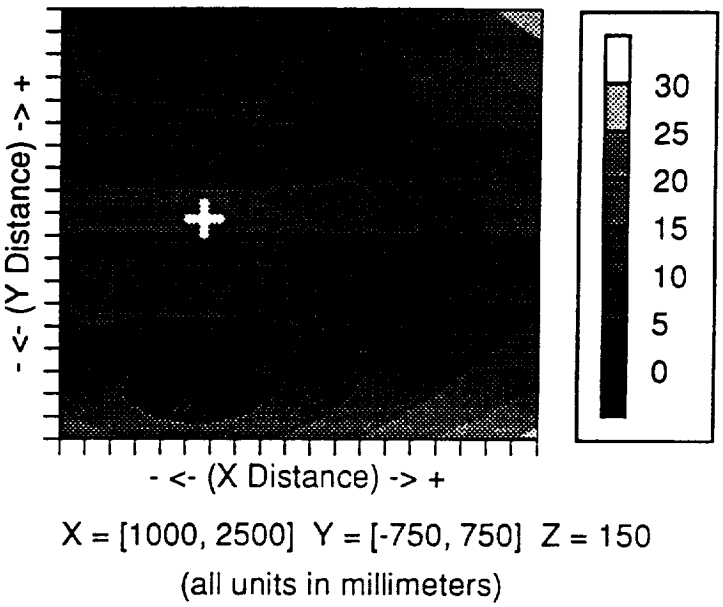


Figure 3.9: Contour plot of the magnitude error caused by perturbing the α parameters. z = 150mm

shift in one axis only. Hence, all points in space should be shifted in the same direction by the same amount.

- Most of the error observed in the simulations results was in the z component of the error vector $P_w\vec{Q}$. This phenomenon is due to the position of the laser. The laser is situated at 2600mm above the testbed. Therefore, the slope of the laser beam as it projects into the workspace will have a large z component, and as a result, a small movement in x or y will result in a large movement in z.

- A 100% increase in the distance between the laser's scanning mirrors (e) had little ($\leq$ .1mm) effect on the accuracy of the laser in comparison to changes in other calibration parameters. This result is due to the fact that the distances from the laser scanner to the planes of testpoints range from 2000mm to

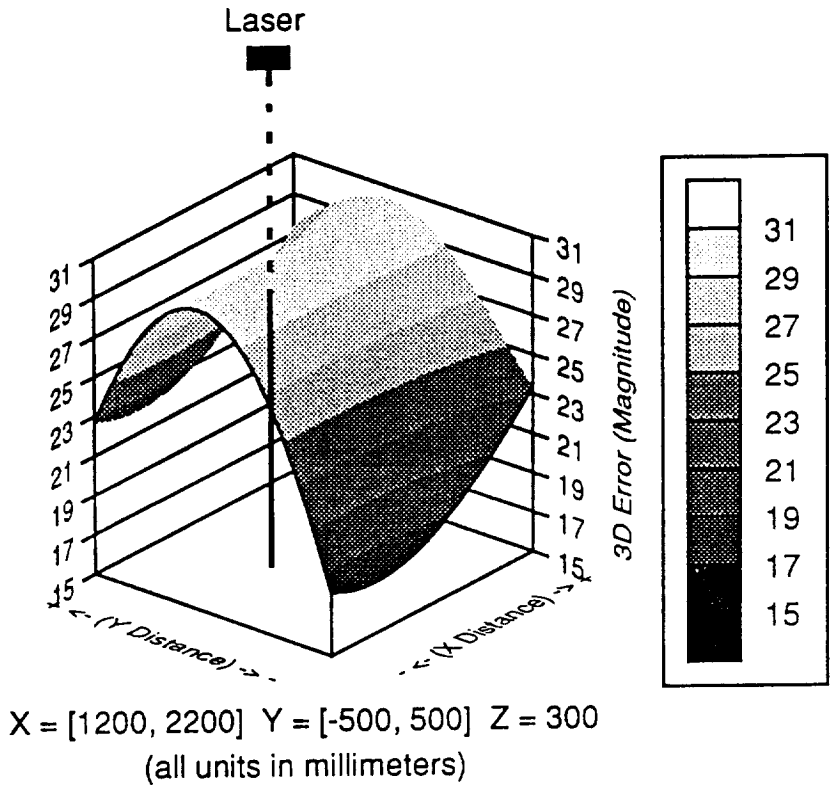


Figure 3.10: Surface plot of the magnitude error caused by perturbing the α parameter. z = 300mm

2600mm. The five millimeter distance between the scanning mirrors is small by comparison to these distances.

- The net effect that an error in a calibration parameter will have on overall accuracy depends on the amount of error and the degree to which the parameter contribute to the pose of the laser scanner. For example, if θ is small, doubling the value of the dx parameter (which influences the value of θ) will have less effect on the laser scanner's pose than if θ were large.

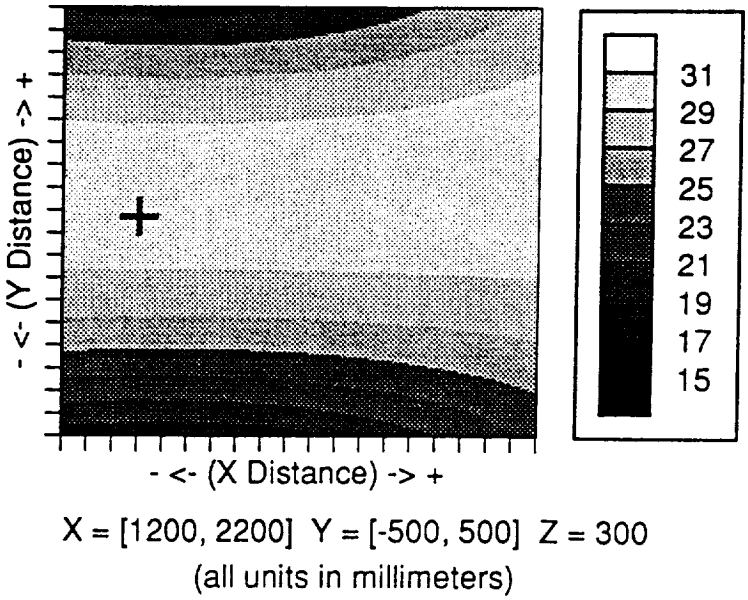


Figure 3.11: Contour plot of the magnitude error caused by perturbing the α parameter. z = 300mm

3.3.3 Simulation of calibrated laser and camera point estimation

A second simulation developed for this research simulates the effect that rotational or translational perturbations will have on point estimates using a calibrated laser and camera. In this case, rotational or translational perturbations are applied to the ${}^{l}_{w}T$ or ${}^{c}_{w}T$ directly. This makes it possible to directly alter the pose of the laser scanner or the camera by rotating or translating about the laser's coordinate frame or about the camera's coordinate frame respectively.

The camera model used in this simulation is the common pin-hole camera model[17]. While other models exist that correct for lens distortion[12] these models adjust pixel coordinates to generate undistorted pixels which can then be applied to the pin-hole camera model. Indeed, the point estimation methods used in this research assume either a pin-hole camera model or that the cameras pixel values have been adjusted for lens distortion so that the pin-hole model can be applied.

It should be noted that the simulation will affect the extrinsic parameters of the camera (i.e. its pose) and these parameters are applied to the camera model *after* any corrections for lens distortion have been performed.

The measure of accuracy used in this simulation is the vector $\vec{P_w P_e}$ from the desired world point to the estimated world point. This measure can provide the absolute error of the estimate by determining the magnitude of $\vec{P_w P_e}$ or it can provide directional information by examining the x, y, z components of $\vec{P_w P_e}$.

To perform an analysis, the simulation performs the following steps:

1. Define a matrix of test points defined in world coordinates. This matrix should define some plane in the workspace.

2. Using ${}^{l}_{w}T$ and equation (2.4) calculate the mirrors angles required to direct the laser beam to each of the points in the matrix.

3. Using ${}^{c}_{w}T$ and the pin-hole camera model calculate the pixel coordinates corresponding to the projection of each point in the matrix onto the camera's image plane.

4. Apply rotation or translation perturbations to ${}^{l}_{w}T$ and ${}^{c}_{w}T$ to obtain ${}^{l}_{w}T'$ and ${}^{c}_{w}T'$. These perturbations are defined in terms of the laser and camera coordinate systems respectively.

5. For each pair of mirror angles and pixel coordinates previously calculated, apply ${}^{l}_{w}T'$, ${}^{c}_{w}T'$ and either the LSE or midpoint to the common normal point estimation method to obtain P_e.

6. Calculate $\vec{P_w P_e}$ for each P_w and corresponding P_e in the matrix.

The definition of the plane of test points is the same as in the laser calibration

simulation. The laser / camera simulation can also generate results in either Matlab, Delta Graph or PvWave formats.

3.3.4 Results of laser and camera simulation

The laser and camera point estimation simulation was subjected to a battery of twelve tests. Each test perturbed one or more rotation and translation parameters of either the laser or the camera. Table 3.4 indicates which parameter was perturbed for each test.

Test	1	2	3
Parameter	θ_l	ψ_l	ϕ_l
Amount changed	0.3rad	0.3rad	.03rad
Test	4	5	6
Parameter	θ_c	ψ_c	ϕ_c
Amount changed	0.3rad	0.3rad	.03rad
Test	7	8	9
Parameter	${}^lt_x$	${}^lt_z$	${}^ct_x$
Amount changed	5mm	-10mm	5mm
Test	10	11	12
Parameter	${}^ct_z$	$\theta_l, \psi_l, \phi_l, \theta_c, \psi_c, \phi_c$	${}^lt_x, {}^lt_z, {}^lt_z, {}^ct_x, {}^ct_z, {}^ct_z$
Amount changed	-10mm	.01rad each	.01rad each

Table 3.4: Parameters perturbed for laser and camera simulation tests

Each test was conducted at three different planes in the CIRSSE testbed. The dimensions of these planes are listed in table 3.5. Each plane had 400 test points distributed evenly over it as a 20 × 20 grid. The coordinates of the planes were chosen to be approximately symmetrical in the x and y axes about the origin of the laser scanner. For each test at each plane, the x, y, z, and magnitude of $\vec{P_uP_e}$ was recorded. Additionally, the LSE and midpoint to the common normal point estimation methods were both applied to each test at each plane. Since the tests generated approximately 400 error plots, the more interesting results are presented in this report while general observations are presented about the remainder of the

tests.

Plane	Min X	Min Y	Max X	Max Y	Z
1	400mm	-1000mm	2400mm	1000mm	0mm
2	400mm	-1000mm	2400mm	1000mm	300mm
3	900mm	-500mm	1800mm	500mm	600mm

Table 3.5: Dimensions of point planes (in world coordinates)

Test seven simulated a translational perturbation along the laser scanner's x axis. The three dimensional surface plot of the error induced by this perturbation at $z = 300$mm is presented in figure 3.12. The projection of the origins of the laser scanner and camera onto the world xy plane are indicated in that figure. Similar results were obtained at $z = 0$mm and $z = 600$mm. The error curve is approximately symmetrical about the laser's origin. In this particular case, a single perturbation had a simple effect on the performance of the point estimation.

When the camera was rotationally perturbed about its y axis, the resulting error curve was particularly interesting. This error curves are depicted in figures 3.13 and 3.14. As is evident from the figures, the accuracy of the point estimates is highly irregular across the plane of test points. Additionally, the shape of the error curve was not consistent across the different planes of test points. At $z = 0$mm (see figure 3.13) the curve achieves a minimum approximately below the origin of the laser, yet at $z = 300$mm (see figure 3.14) the error curve reaches a maximum at the same location. The reason for such unusual behavior is similar to the reason for the behavior observed when the α parameter was perturbed in section 3.3.2 in that individual terms are competing for "dominance" at different regions of the workspace.

The results of the tests conducted using the laser and camera simulation produced several observations about the performance of the LSE and midpoint to the

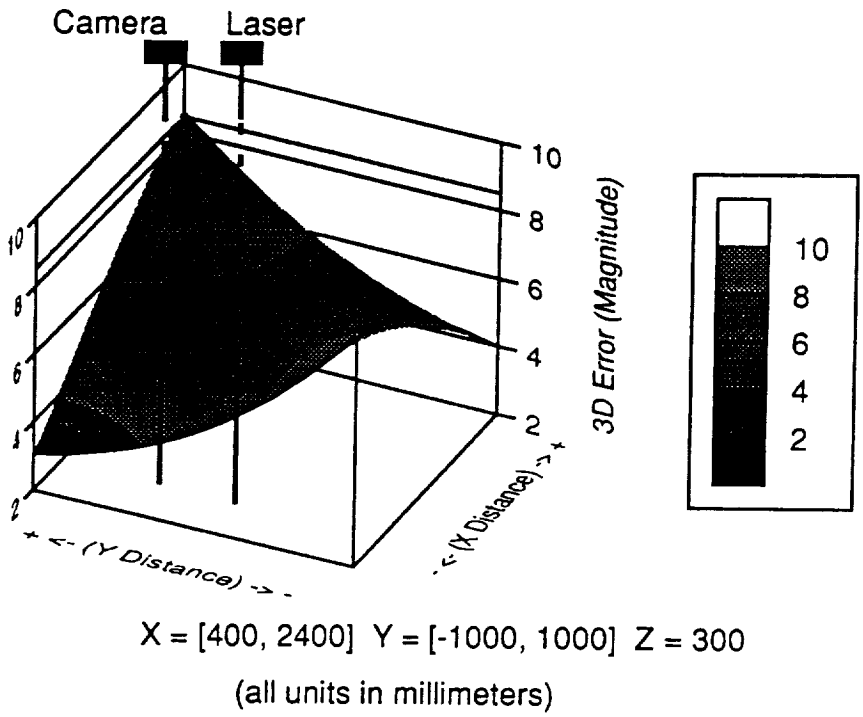


Figure 3.12: Surface plot of the magnitude of error caused by perturbing the camera along its x axis. $z = 300$mm

common normal point estimation techniques:

- The estimates generated by the midpoint to the common normal and the LSE techniques usually were identical. In those cases where the estimates were different, the midpoint to the common normal *usually* exhibited less error, however there were cases where the LSE produced better estimates. It is important to remember that the LSE method assumes that optimizing a system of linear equations will minimize the point estimation error, while the midpoint to the common normal method assumes that the midpoint of the normal between two rays is the best point estimate. These assumptions are reasonable, but there will be cases of laser and camera positions and point locations where one assumption is better than the other.

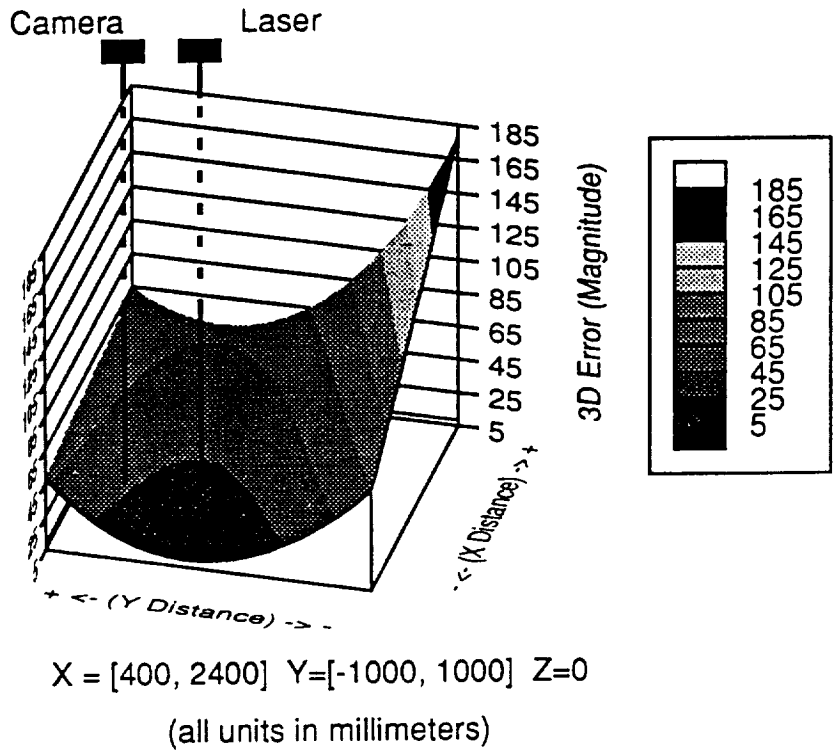


Figure 3.13: Surface plot of the magnitude of error caused by perturbing the camera's θ Euler angle. z = 0mm

- Rotational perturbations about either the camera or laser produced errors that were highly irregular over a given plane and were inconsistent across different planes. The reason for this is similar to the same observation made in section 3.3.2. A rotation shift affects two axes of the coordinate system and does so as a nonlinear function. A translational shift affects one axis by a constant value.

- Translational perturbations generally produced errors that were symmetrical for a given plane and consistent across different planes. This is again due to the fact that a translational shift will affect only one axis by a constant amount. Such a constant shift would be visible and constant throughout the

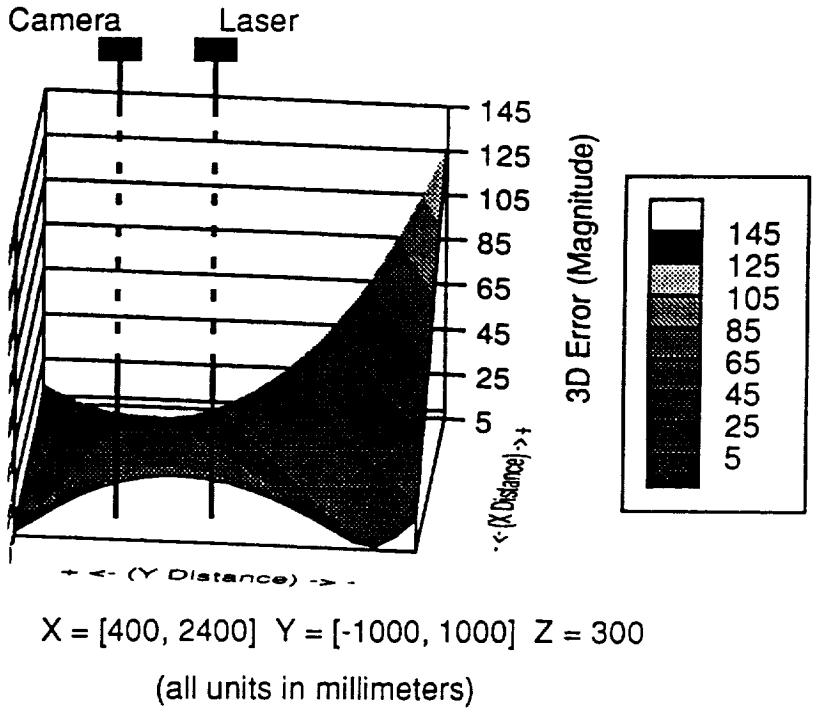


Figure 3.14: Surface plot of the magnitude of error caused by perturbing the camera's θ Euler angle. z = 300mm

entire workspace.

- Most of the error in the point estimates was in the z component of $\vec{P_w P_e}$. This is due to the physical locations of the laser and camera. These devices are situated 2600mm above the testbed floor. Hence, rays projecting from these devices in to the testbed will have slopes with large z components. As a result, a small movement in x or y will result in a large movement in z. If the laser and camera were placed further apart, the z component of the error would be less because the ratio between the z components of the slope and the x and y components would be reduced.

The results of these tests indicate that the simulation can be a useful tool in applying a laser scanner and camera to three dimensional sensing. If the errors in

the pose of the laser scanner and camera can be identified, this simulation could be used to determine which estimation method should be used in different parts of the workspace to achieve the best results.

It is important to understand why the point estimation algorithms exhibited such unusual behavior. In these tests the accuracy of the point estimates are dependent on the degree of skew between the projected rays from the laser and camera and the desired point. The amount of skew between the rays is dependent on the pose of the laser scanner and camera, the errors applied to the their pose, and the coordinates of the test point. In some cases, the projected rays may pass at equal distances from the desired point, in which case their respective errors might cancel out yielding an accurate estimate. In other cases, however, the one ray may be closer to the desired point. Hence, the estimate would be shifted toward the ray that is further away from the desired point. Since there are so many variables involved, it is not surprising that the error curves obtained in figures 3.14 - 3.13 would be observed.

3.4 Locating a Laser Spot in a Camera Image

To use a calibrated laser scanner in concert with a camera, it is necessary to be able to locate the laser spot in the camera image. This is a simple problem if one can guarantee that the laser spot is the brightest region in the image. However, such an assumption restricts the utility of a calibrated laser by placing illumination constraints on the image. If techniques can be employed to locate the laser spot in the presence of "noise" (e.g. pixels of similar intensity), then a calibrated laser can be used in a wider variety of situations. The method developed by the author to locate the laser spot in a noisy image is a heuristic approach whereby regions in the image are successively eliminated based on a set of criterion tests.

The first step in locating the laser spot is to perform region growing[18] over some selected area of the image. This results in a list of regions, their area, and their centroids. The laser spot should be one of the regions in this list. To isolate the laser spot it is necessary to eliminate all those regions that are not attributable to the laser. There are four different tests that can be applied to the region list to perform this elimination. Each test returns a list of the regions that passed the test. The laser spot should be the only region that passes all four tests.

The first test is to eliminate all regions that do not fall within a specified intensity range. Since the laser will appear as a small bright spot, it will be one of the brighter regions in the image. However, there is no guarantee that the laser spot will be among the brightest. Indeed, if the laser beam is illuminating a matte (low reflectivity) object, such as a piece of cloth, then the intensity of the laser spot will be lower than if the beam was reflecting off a piece of metal. Additionally, specular reflections of ambient light off high reflectivity objects can exhibit the same intensity as the laser spot.

A second test that can be applied to the region list is to eliminate regions that do not fall within a certain range of sizes. The laser spot typically occupies between two pixels and twenty pixels depending on the reflectance of the object the laser beam is striking. The tests for size and intensity can detect the presence of a laser spot in the image in most cases. Problems arise when there are other regions in the image that have the same size and intensity characteristics as the laser, such as specular reflections.

If after application of the intensity and size tests the region list still has more than one candidate region, two more tests can be applied to further reduce the list. One of these tests is to take the centroid coordinates of each region and the known scanning mirror angles, run them through a point estimation algorithm, and

eliminate the regions that generate solutions that fall outside of the workspace. This method does assume that some *a priori* knowledge exists about the expected location of the laser spot. The more that is known about the expected location of the laser spot, the greater the chance of correctly identifying it. Usually, little *a priori* knowledge is required to locate the laser spot, since only those regions that lie along the laser ray will generate results that are reasonable.

A final test can be employed if all the previous tests have failed to return a unique solution for the laser spot. The laser beam can be moved and another image acquired. The new image is passed through the region growing algorithm just as the first image. If the scene is static, the only region that should have moved is the laser spot.

It is important to note that it is not necessary to use all four tests. If a subset of tests yields one region, then the remaining tests do not have to be run. Further, there are situations where it may be impossible to locate the laser spot. If the laser spot is within the bounds of a bright region, the camera may not be able to distinguish it. This problem is particularly acute if the camera's aperture is too wide, since bright regions could then saturate the camera's CCD element. The laser spot is also undetectable if it is physically occluded by an object in the workspace. It should be noted that these four tests do not necessarily have to be performed in the order stated above. Indeed, part of the evaluations presented later in this chapter address the question of an optimal ordering for these tests.

3.4.1 Application of region growing algorithm to a camera image

Once a camera image is acquired, a region of interest is selected whose boundaries are such that it encompasses the laser region. The pixels within the region of interest are grouped into regions of similar intensity. The algorithm employed

in this research is similar to the blob coloring algorithm proposed by Ballard and Brown [18]. The specific heuristic algorithm used in this application is described in figure 3.15.

A pixel is considered part of a unique region if its intensity is similar (by T_{region}) to its top, left, or top-left neighbors. The algorithm generates a list of regions identified in the image. Each entry in the list contains data on the regions size (in pixels), maximum intensity, centroid, and equivalence to another region in the list. The concept of region equivalency deserves more explanation. Envision performing this algorithm on an image that contained a region that is shaped like the letter "U". As the pixels are scanned the top portion of the "U" would be identified as two distinct vertical regions. At the point where the pixels form the curve at the bottom of the "U", the algorithm will find that the two regions its has been growing are actually the same region. In this case, an equivalency pointer in one region is set to the value of the other region.

The number of regions that are generated by this algorithm depends on the values of T_{back} and T_{region}. T_{back} essentially dictates how much of the image is eligible for region growing; while T_{region} determines how much contrast is required between pixels before a new region is detected. While the specific values for these variables depends on lighting conditions and image complexity, the values used in the CIRSSE testbed for room lighting conditions are typically $\{150 \leq T_{back} \leq 190\}$, $\{8 \leq T_{region} \leq 10\}$.

After the region list is constructed, it is assumed that any of the regions could be the laser spot. The next step after region growing is to eliminate from consideration all those regions that are equivalent to other regions (i.e. those regions that do not have the equivalency pointer equal to itself). At this point, the region list contains N unique regions.

for each pixel **from** left to right **and** top to bottom **do**:

if pixel intensity > background threshold T_{back}

if pixel directly above current pixel is part of a region **and** intensity of current pixel is similar to pixel directly above to within a given threshold T_{region}

- Mark current pixel as belonging to the same region as the pixel directly above it

else if pixel to left is part of a region **and** intensity of current pixel is similar to pixel directly above to within a given threshold T_{region}

- Mark current pixel as belonging to the same region as the pixel directly above it

else if pixel to the top left is part of a region **and** intensity of current pixel is similar to pixel directly above to within a given threshold T_{region}

- Mark current pixel as belonging to the same region as the top-left pixel

else

- Current pixel is part of a new region

else

pixel is part of the background

/* Check for region equivalence */

if {pixel to left of current pixel is part of a region} **and** {pixel above is part of a different region} and {the intensities of the two regions are similar to within T_{region}}

- Region to left of current pixel is equivalent to the region above the current pixel.

end loop

Figure 3.15: Region growing algorithm

3.4.2 Region selection based on intensity and size

Regions can be eliminated based on intensity and size criterion. This process is straightforward in that each region in the list is checked to see if it falls within a range of intensity values or within a range of size values depending on which criteria is being used. To ensure efficient performance, the size and intensity tests should be performed only on those regions that have not been eliminated as possible laser regions.

The execution time of the intensity algorithm can be expressed as:

$$\begin{aligned} \Gamma(intensity) &= C_{valid}(N) + C_{intensity}(n_{l1}) \\ &\quad \{0 \leq n_{l1} \leq N\} \end{aligned} \tag{3.28}$$

Where C_{valid} represents the execution time required to determine if a region is a possible laser region and $C_{intensity}$ represents the time required to determine if the region's intensity falls within the specified limits. Since both of these operations consist of if/then comparisons the computation times for these operations can be expressed as constants. n_{l1} represents the number of regions in the list that have not been eliminated as laser regions.

Similarly, the execution time to eliminate regions by size can be expressed as:

$$\begin{aligned} \Gamma(size) &= C_{valid}(N) + C_{size}(n_{l2}) \\ &\quad \{0 \leq n_{l2} \leq N\} \end{aligned} \tag{3.29}$$

Where C_{size} is the execution time to determine if a region's size falls within specified limits, and this value is also a constant. In practice, C_{size} and $C_{intensity}$ are

approximately equal because the actual operation being executed in both cases is a simple comparison of integer values. As a result, $\Gamma(intensity)$ and $\Gamma(size)$ are also approximately equal. The execution time analysis for these two algorithms is trivial, yet, as will be discussed later, the performance of these algorithms is critical in determining the order of execution for all the selection criteria.

3.4.3 Region selection based on laser/camera triangulation

Another method for determining the laser region is to apply the three dimensional point estimation algorithms specified in chapter 3.2 to each region in the region list. Assuming the deflection of the laser scanner's mirrors is known, and the laser and camera are calibrated, the centroid data from each region can be used to estimate the three dimensional location of each region. If there is some knowledge as to the three dimensional location of the laser spot, the three dimensional point estimates can be successively eliminated until only those points that are consistent with the expected value remain.

The primary issue at hand is how much knowledge of the three dimensional location of the laser spot is needed to yield a unique solution. In practice, the laser region must lie along the projection of the laser ray across the camera's image plane. If the centroid of a region deviates from this projected line, then the point estimation algorithm will be trying to triangulate two divergent rays.

In practice, the three dimensional point estimates for regions other than the laser spot become highly irregular and minimal knowledge of the laser spot location is needed to reduce the set of point estimates to a unique solution. For example, in the CIRSSE testbed, the world origin is located about 10cm above the floor with the Z axis directed up at the ceiling. If a point estimate yields a Z of (-80cm), this implies that the laser spot is located somewhere in or under the concrete floor of

the testbed, and such a condition is clearly impossible.

The implementation of the triangulation algorithm currently used by the author employs the LSE point estimation algorithm described in section 3.2.1. Each region that has not been eliminated as a possible laser region is passed through the point estimation algorithm and the region is either eliminated or accepted if the estimated point lies within a specific three dimensional volume. Typically, this volume is centered about an estimated position of the laser spot and is constrained to +/- (5-10cm) in each axis about this position.

The execution time of this algorithm can be expressed as:

$$\begin{aligned} \Gamma(triangle) &= C_{valid}(N) + C_{triangle} n_{l3} \\ & \{0 \leq n_{l3} \leq N\} \end{aligned} \tag{3.30}$$

Where n_{l3} is the number of regions in the region list that are possible laser regions and $C_{triangle}$ is a constant representing the computation time to estimate a three dimensional point for an arbitrary region. The $C_{triangle}$ term deserves more explanation. The LSE point estimation algorithm uses singular value decomposition to calculate the estimated point. The execution time of the algorithm is dependent on the size of the A and b matrices. The size of these matrices is dependent on the number of sensing devices used to estimate the three dimensional point. In the case of one laser and a camera, the A matrix is four rows by three columns wide and the b matrix is four rows by one column. Since the number and type of sensing devices *should* not change during the middle of the triangulation algorithm, the dimensions of the A and b matrices will not change. Hence, the execution time for the singular value decomposition algorithm will be the same for each estimated point, and this value can be expressed as a constant. It is also important to mention that the

for each region in current region list **do**

if current region X has not been eliminated as a candidate laser region

for ith region $\{0 \leq i \leq N\}$ in previous region list **do**

if distance between centroid of region X and region i $\leq$ $MaxDistance$ **and** difference between size of region X and size of region i $\leq MaxSizeDiff$ **and** difference between intensity of region X and intensity of region i $\leq MaxIntensityDiff$

- Eliminate region X as a possible laser region

Figure 3.16: Algorithm for elimination of regions based on movement

value of $C_{triangle}$ is *much* larger than either $C_{intensity}$ or C_{size}. Hence, while all three algorithms execute in $O(N)$ time, the triangulation algorithm requires greater time to execute that the size or intensity algorithms.

3.4.4 Region selection based on movement

Another method of determining the laser region is to acquire one image, move the laser, acquire a second image, and then eliminate all those regions that did not move. The algorithm used to determine if a region has moved is presented in figure 3.16.

The algorithm in figure 3.16 will identify new regions in the current region list as regions that have moved. This is due to the fact that a new region in the current list cannot be correlated to a region in the previous list. This characteristic of the algorithm is neither a drawback nor an advantage as much as it is necessary to understand that the algorithm behaves in such a manner. The execution time of the algorithm can be expressed as:

$$\Gamma(movement) = C_{valid}N_{current} + C_{movement}n_{14}N_{previous}$$

$$\{0 \leq n_{l4} \leq N_{current}\} \tag{3.31}$$

As one might expect, $N_{current}$ and $N_{previous}$ are the number of regions detected in the current image and the previous image respectively, and $C_{movement}$ is the execution time required to determine if a single region has moved from the previous frame. Assuming that $N_{current} \approx N_{previous}$ the algorithm in figure 3.16 executes as $O({N_{current}}^2)$ in the worst case.

3.4.5 Evaluation of laser region identification performance

A treatment of the issue of detecting a laser spot in a camera image would not be complete without a thorough evaluation of the performance of the algorithms under experimental conditions. The region selection algorithms were combined into a single program that directs the laser to specific three dimensional points and subsequently acquires images of the workspace for each point using a camera. This program was subjected to four different batteries of tests to determine the behavior of the laser spot selection algorithms to varying experimental conditions. The descriptions of the four test batteries are presented below:

- *Test Battery 1*: While maintaining constant lighting and region detection parameters, vary the complexity of the image by adding objects of differing size, and reflectance qualities. Low complexity images had few objects such as cables, a few bits of metal and so forth, while more complex images contained everything in the low complexity images plus struts and unpainted metal nodes. Struts are metal rods approximately 60cm long and 2cm wide while nodes are hexagonal structures approximately eight centimeters wide. Struts are linked together with nodes to build structures.

- *Test Battery 2*: While maintaining constant scene complexity and region detection parameters, vary the location and intensity of scene illumination.

- *Test Battery 3*: While maintaining constant scene complexity and illumination, place objects over a wide range of three dimensional locations in the workspace. The objective here is to determine how the triangulation algorithm performs when the valid three dimensional volume is set to encompass a large portion of the workspace.

- *Test Battery 4*: While maintaining constant scene complexity and illumination, vary the order in which the region selection algorithms are executed.

The parameters used for the region selection algorithms are summarized in table 3.6. The X, Y, and Z parameters for the triangulation algorithm define the valid three dimensional volume for region point estimates. The volume is defined with respect to the X, Y, and Z coordinates of the expected location of the laser spot. The expected location of the laser is determined by instructing the program to direct the laser beam at a specific world point. The three dimensional volume was changed for test battery three since the objective of these tests is to determine if the triangulation algorithm will work with a large valid volume.

The results of each test are encapsulated in eleven parameters defined as follows:

- *AvgRegion* The average number of regions detected by the region growing algorithm over all test trials.

- *AvgEquiv* The average number of regions eliminated as the laser region due to equivalency over all test trials.

Intensity	Min	210
	Max	255
Size	Min	4 pixels
	Max	50 pixels
Movement	MaxDistanceDiff	5 pixels
	MaxIntensityDiff	15
	MaxSizeDiff	10 pixels
Triangulation Tests 1,2,4	$\{-50 \leq X \leq 50\}$ mm $\{-50 \leq Y \leq 50\}$ mm $\{-50 \leq Z \leq 100\}$ mm	
Test 3	$\{-200 \leq X \leq 200\}$ mm $\{-200 \leq Y \leq 200\}$ mm $\{-350 \leq Z \leq 500\}$ mm	

Table 3.6: Parameters used for test batteries

- *AvgIntensity* The average number of regions eliminated as the laser region due to selection by intensity over all test trials.
- *AvgSize* The average number of regions eliminated as the laser region due to selection by size over all test trials.
- *AvgMovement* The average number of regions eliminated as the laser region due to selection by movement over all test trials.
- *AvgTriangle* The average number of regions eliminated as the laser region due to selection by triangulation over all test trials.
- *NoLaser* The number of test trials where no laser region was found.
- *Laser1* The number of test trials where one laser region was found.
- *Laser2* The number of test trials where two laser regions were found.
- *Laser3* The number of test trials where three laser regions were found.

Parameter	Test 1	Test 2	Test 3
AvgRegion	96	129	546
AvgEquiv	15	17	89
AvgIntensity	36	52	200
AvgSize	33	45	203
AvgMovement	10	12	49
AvgTriangle	0/1	0/1	2
NoLaser	3	4	8
Laser1	97	96	90
Laser2	0	0	2
Laser3	0	0	0
LaserGT3	0	0	0
Image Complexity	Low	Moderate	High

Table 3.7: Results of test battery one

- *LaserGT3* The number of test trials where more than three laser regions were found.

Each test in each test battery consisted of one hundred laser points. T_{back} and T_{region} were set to 190 and 10 respectively.

3.4.5.1 Analysis of test battery 1 results

The first battery of tests was designed to study the behavior of the laser spot selection algorithms to varying scene complexity. The first test was conducted on a scene of low complexity that contained a few metal objects and a cable. The second test was conducted on a scene of higher complexity that included more multi-faceted metal objects and a few hand tools. Finally, the third test was conducted on a highly complex scene that included struts and nodes (previously defined), grippers from the CIRSSE robot arms and a metal plate. The results of each test are presented in table 3.7.

The results of the three experiments indicates that elimination by size and by intensity had the greatest effect in reducing the number of regions in the image.

Further, the triangulation algorithm was not required for the first two tests, which implies that size, intensity, and movement are sufficient criteria for identifying the laser in simple or moderately complex images. What is most significant about the test results is the number of times the laser spot was identified. The worst case results indicate that the selection algorithms produced a unique solution for the laser spot 90% of the time. What is even more interesting, is that there are much fewer instances of identifying more than one region as the laser. Indeed, the results indicate that the selection algorithms either found a single laser spot or none at all.

The instances where the laser was not found can be attributed, in most cases, to effects of illumination or occlusion that prevented the camera from distinguishing the laser region. For example, it was observed in several instances that the laser spot was projected near an object such that the object occluded the laser spot from the camera's field of view. Further, there were regions in the image that registered as full intensity (i.e. 255) and when the laser was directed into these regions, the camera was unable to distinguish the laser spot since it was embedded in a region that saturated the CCD array. In short, failure to detect the laser region altogether is due mostly to the inherent limitations of the equipment used to perform the test and not to the performance of the algorithms.

3.4.5.2 Analysis of test battery 2 results

The second test battery was designed to study the effects of illumination on the selection algorithms' performance. The first test in this set was performed on an image of moderate complexity (similar to test 2 in the first test battery) under normal room lighting conditions. This first test provided a baseline for gauging performance of other tests. The second test was conducted with the lights off. The third test was conducted with a single light source projected from one end of the

scene (the top of the image) and oriented to place strong shadows on the objects in the scene. In the final test, a single light source was oriented not only to project strong shadows on the objects in the scene, but also to project reflection spots and halo effects into the camera's lens. The results of these tests are presented in table 3.8.

Parameter	Test 1	Test 2	Test 3	Test 4
AvgRegion	474	3	368	568
AvgEquiv	63	0/1	46	131
AvgIntensity	213	0/1	154	176
AvgSize	159	1	127	227
AvgMovement	31	0/1	37	28
AvgTriangle	5	0/1	1	3
NoLaser	5	1	5	15
Laser1	93	99	94	84
Laser2	2	0	1	1
Laser3	0	0	0	0
LaserGT3	0	0	0	0
Illumination	Normal	No lights	Shadow	Shadow & lens reflection

Table 3.8: Results of test battery two

The results indicate that for the first three tests, the selection algorithms were largely resilient to changes in ambient light, in that the selection algorithms achieved a unique solution in more that 93% of the trials. What is particularly interesting is how performance degraded in test 4. It is not surprising that performance would degrade if light is directed in the camera, but what is interesting is that the pattern of performance is nearly identical to the results obtained in the most complex image of the first test battery. While the similarity in the exact numbers may be a coincidence, the pattern indicates that the performance of the selection algorithms degrades in a consistent manner. That is, the algorithms either achieve a unique solution for the laser spot or none at all. When test 4 was actually conducted, it was observed that the laser region was lost when the laser was directed into a

region of the scene that was highly illuminated (to the point where the camera was saturated) or the laser was directed into a region of the image that contained a lens reflection. In both of these cases, the intensity of the image registered as 255 which is the maximum intensity value for the vision system, and therefore, the laser spot was visually indistinguishable from the surrounding image.

3.4.5.3 Analysis of test battery 3 results

The third test battery was designed to test the effectiveness of the triangulation algorithm if the valid three dimensional volume was enlarged to cover a greater portion of the testbed. In this test, objects were placed in the workspace in such a way as to ensure that specular reflections and other noise were present over a wide three dimensional volume in the workspace. This arrangement was adopted to increase the chance that specular reflections would lie along the laser ray thereby increasing the probability that the selection algorithms would misidentify some of these regions as being attributable to the laser. The results for test battery three are presented in table 3.9. The two tests were virtually identical, although the arrangement of objects in the scene was altered between tests to provide different scenes of similar complexity.

There are two significant observations that can be made about these results. First, while enlarging the valid three dimensional volume does increase the probability that more than one region in an image will be identified as the laser spot, in the large majority of cases the selection algorithms either achieved a unique solution for the laser spot or could not find the laser region at all. In other words, the general behavior of the selection algorithms in this battery of tests is similar to the the other test batteries.

A second observation was made while studying the behavior of the selection

Parameter	Test 1	Test 2
AvgRegion	499	189
AvgEquiv	103	41
AvgIntensity	208	74
AvgSize	132	52
AvgMovement	43	17
AvgTriangle	9	2
NoLaser	17	8
Laser1	68	79
Laser2	5	11
Laser3	4	2
LaserGT3	6	0

Table 3.9: Results of test battery three

algorithms as the experiments were conducted. In cases where the selection algorithm generated multiple solutions for the laser spot the regions in question were either in close proximity to the actual laser spot (to within a few millimeters) or were a significant distance away from the actual laser spot and situated along the projection of the laser ray through the image. These observations are not surprising when one considers that the laser ray and the projected ray from the camera to these regions are not significantly divergent. To eliminate regions that do not result in divergent rays, it is necessary to be able to make a more accurate estimate of where the laser spot is expected. However, as is apparent by these test results, even a rough *a priori* estimate of the expected location of the point usually results in selecting the correct region as the laser spot.

3.4.5.4 Analysis of test battery 4 results

The fourth test battery was designed to determine if changing the execution order of the selection algorithms resulted in a significant change in performance. The results of these tests are presented in table 3.10.

The results of these tests indicate that while the pattern of region elimination

Parameter	Test 1	Test 2	Test 3
AvgRegion	362	363	364
AvgEquiv	38	38	38
AvgIntensity	115	6	6
AvgSize	174	286	15
AvgMovement	30	29	306
AvgTriangle	1	1	1
NoLaser	6	5	5
Laser1	92	94	93
Laser2	2	1	2
Laser3	0/1	0/1	0
LaserGT3	0/1	0/1	0
Order of Execution	Intensity Size Movement Triangle	Size Intensity Movement Triangle	Movement Size Intensity Triangle

Table 3.10: Results of test battery four

differs between the different orders of algorithm execution, the final results for identifying the laser spot are virtually identical across all the tests. The implications of this result is that the order of execution can be arranged to optimize the overall performance of the selection algorithms without sacrificing reliability.

To optimize performance of the selection algorithms it is necessary to examine the equations for execution time of each algorithm that were derived previously. The total execution time of the laser selection process is the sum of these individual equations as indicated below:

$$\begin{aligned} \Gamma(total) &= \Gamma(intensity) + \Gamma(size) + \\ &\quad \Gamma(movement) + \Gamma(triangle) \qquad (3.32) \\ \Gamma(intensity) &= C_{valid}(N) + C_{intensity}(n_{l1}) \qquad (3.33) \\ \Gamma(size) &= C_{valid}(N) + C_{size}(n_{l2}) \qquad (3.34) \end{aligned}$$

$$\Gamma(movement) = C_{valid}N_{current} + C_{movement}n_{l4}N_{previous} \quad (3.35)$$

$$\Gamma(triangle) = C_{valid}(N) + C_{triangle}n_{l3} \quad (3.36)$$

where the expressions for $\Gamma(intensity)$, $\Gamma(size)$, $\Gamma(movement)$, and $\Gamma(triangle)$ are restated here for convenience. The total execution time is dependent on the total number of regions in the image (N) and the number of possible laser regions passed to the individual selection algorithms (n_{l1}, n_{l2}, n_{l3}, n_{l4}). Recall from the previous discussion that $\Gamma(movement)$ executes in $O(N^2)$ time in the worst case and $C_{triangle} \gg \{C_{intensity}, C_{size}\}$. Therefore, the best way to reduce overall execution time is to reduce the contribution of the $\Gamma(movement)$ and $\Gamma(triangle)$ terms. This can be accomplished by keeping n_{l3} and n_{l4} small. In other words, use the size and intensity algorithms to eliminate as many regions as possible before executing the more computationally complex movement and triangulation algorithms. Since the execution time for the intensity and size algorithms are roughly equivalent (see section 3.4.2), it does not matter whether the size test is executed before the intensity test or vice versa. What is important is that these two tests should be executed before the movement and triangulation tests so as to minimize n_{l3} and n_{l4}.

At this point, it is clear that the intensity and size algorithms should be executed first to reduce the number of possible laser regions that must be sent through the movement and triangulation algorithms. The next issue is whether the movement algorithm or the triangulation algorithm should be executed next. The results from all the tests indicates that $\{n_{l3} \approx n_{l4}\} \ll N$, and, hence, the movement algorithm's execution time will more closely approximate $O(N)$ instead of $O(N^2)$.

The execution time of the triangulation algorithm also approximates $O(N)$. It was mentioned previously that the value of $C_{triangle}$ is significantly larger than

$C_{intensity}$ and C_{size}. In fact, $C_{triangle}$ is also much larger than than $C_{movement}$ which represents the execution time of a small number of comparisons to determine if two regions are equivalent. When n_{l3} is small, the value of the the execution time for the triangulation algorithm is comparable to that of the movement algorithm despite the fact that the movement algorithm theoretically should be less efficient. Therefore, the total execution time of the laser selection algorithms will be relatively constant regardless of whether movement is executed before triangulation or vice-versa.

3.4.5.5 Conclusions about laser selection performance

The results presented in the previous sections have provided a plethora of information about the behavior of the laser selection algorithms, both individually and in concert with each other. From the results and the subsequent analysis it is possible to draw several conclusions about the performance of these algorithms.

1. The combination of all four selection algorithms locates the laser spot reliably under normal lighting conditions and moderate to high scene complexity.

2. Degradation of algorithm performance results in a decreased potential for locating the laser spot in the image as opposed to inadvertently selecting multiple regions as the laser spot.

3. Algorithm performance is most affected by occlusion of the laser spot in the workspace and saturation of the camera's CCD array due to aperture setting.

4. Laser selection algorithms operate most efficiently when the intensity and size selections are executed first followed by movement and triangulation.

There are other properties of the laser spot that may be useful in enhancing the discrimination of the spot in a camera image. First, the laser spot has a specific

spectral wavelength and if the camera were fitted with a filter that is sensitive only to the wavelength of the laser, it would be easier to locate the laser spot. Additionally, the laser spot has an elliptical or circular geometry. Therefore, if the region growing algorithm were modified to record more information about the geometry of each region such as the length of the region's perimeter and the region's moments of inertia, it might be possible to eliminate regions that do not resemble a small ellipse.

Overall, the techniques discussed in this section for locating a laser spot in an image offer a reliable method for laser tracking under a variety of lighting and scene conditions. These methods will permit a calibrated laser and camera to operate under the same conditions as multiple camera configurations. Such a capability permits a laser and camera to not only be a useful active three dimensional sensing device in its own right, but it also provides the ability to verify results obtained by passive techniques.

3.5 Representation of Three Dimensional Data

The techniques described up to this point optimize the performance of the laser and camera sensing configuration by improving the accuracy and reliability of determining the three dimensional coordinates of the laser spot. Beyond a certain point, however, improving the accuracy will result in a marginal improvement in system performance. The limiting factor will be the amount of three dimensional information that can be extracted from the data points. A collection of data points is more useful if some structure can be associated with the points. Hence, to fully realize the potential of the laser and camera sensing configuration it is necessary to develop techniques to represent and manipulate three dimensional data.

In order to provide three dimensional information about the workspace, it is

advantageous to repeatedly subdivide the workspace into smaller regions. Specifically, the workspace should be divided into a collection of planar surfaces. While any polygon could be used to define these planar surface, the triangle has several qualities that make it the best choice. First, since three points uniquely determine a plane, the triangle is the simplest planar structure. Second, the centroid of a triangle always lies within the plane defined by its vertices. Third, triangles with common edges can be combined to form more complex polygons.

While it is fine and good to decide that the workspace should be divided into triangular surfaces (or facets), the issue at hand is how to generate triangular surfaces from a collection of three dimensional points. Randomly directing the laser beam through the workspace will generate data points, but won't provide the required structure. To provide the planar structure, it is necessary to systematically direct the laser scanner into the workspace and keep track of how the data points relate to each other.

Figure 3.17 depicts an arbitrary triangle with vertices A, B and C. Three new points D E F are defined at the midpoints of the edges AB, BC and AC respectively. These new points can be used in conjunction with the vertices A, B and C to divide the original triangle into four smaller triangles. The edges of each of these smaller triangles are approximately half the length of the edges of the original triangle. If this process of dividing triangles is recursively repeated, the original triangle will be divided into a mosaic of small triangular facets. This approach is similar to methods used in finite element mesh generation [19]. One common technique in finite element mesh generation is to recursively divide the domain into smaller regions[20].

In order to apply the triangle division process to three dimensional sensing, a triangle is defined in some portion of the workspace that is of interest. The vertices

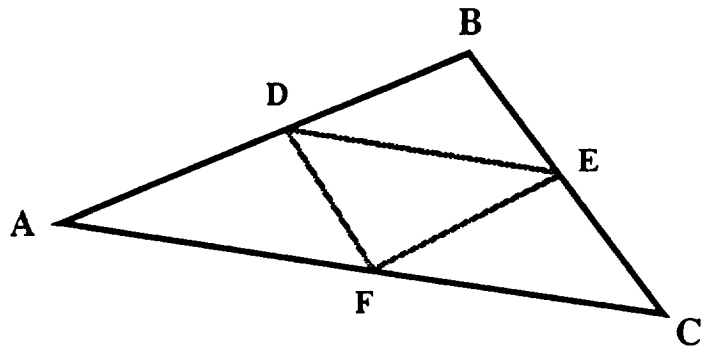


Figure 3.17: Division of a triangular plane

of this triangle are stored in a list of points. The triangle is then divided; thereby creating four smaller triangles and three new points. These four new triangles represent four children of a quad tree whose root is the original triangle. The three new points that are created are then saved into the point list. Each of the smaller triangles is subdivided to create another level in the quad tree and more points in the point list. An example of this process is depicted in figure 3.18. The triangles are recursively divided until the size of the triangles reach some predefined limit. The result of this process is a list a points and a quad tree of triangles, the vertices of which are indexes into the point list. Additionally, the facets can be defined in a counter-clockwise form. This is useful in determining if a collection of facets forms a convex or concave surface.

This method of generating triangular surfaces can be applied in a variety of ways to collect and enhance three dimensional data. The point list can be used to control the laser by directing the beam at each of the points in the list. As the laser beam travels to the intended point and beyond, it will eventually strike an object and generate a laser spot. The camera will detect the laser spot and the "true" three dimensional coordinates of the point will be recorded. These new coordinates

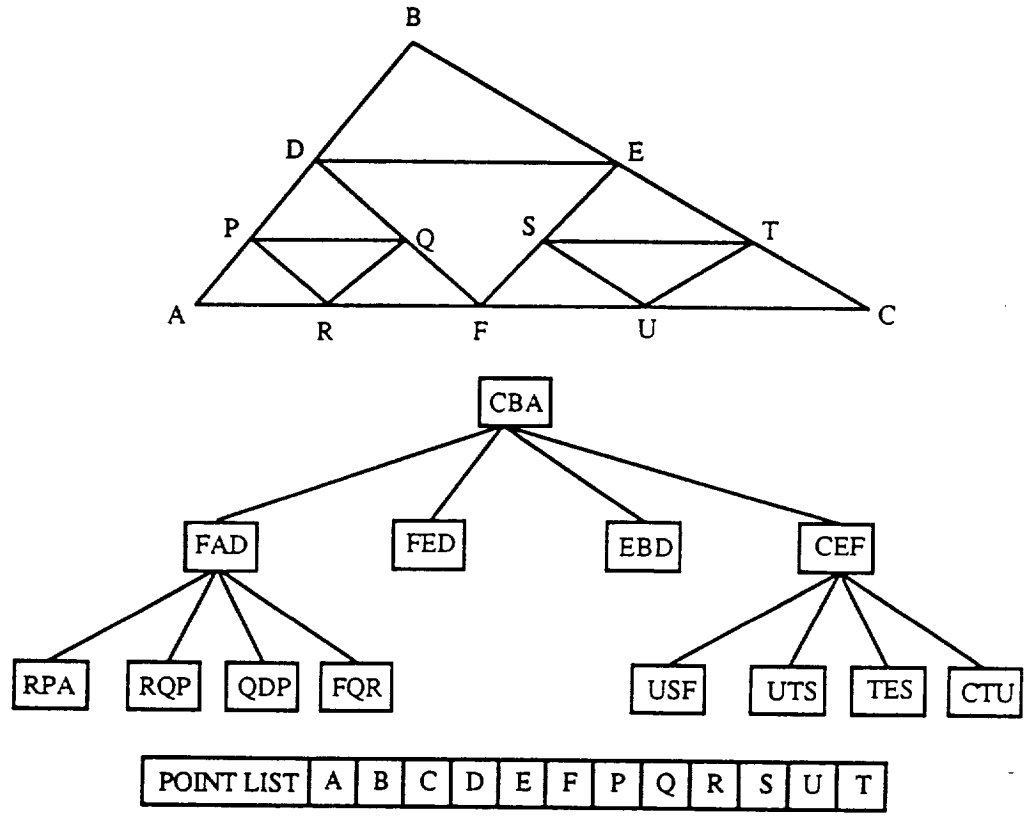


Figure 3.18: Division of workspace into quad tree of triangles

can be stored in a new list and a mapping from the old point list to the new one can be established. At this point a set of data points from the workspace exists along with a quad tree representing how the points relate to each other to generate triangular surfaces.

Another application of the triangular structure is approximating three dimensional points between the vertices of the triangle. For example, suppose it is necessary to estimate the three dimensional coordinates of a point on the triangle but only the x and y coordinates are known. In this case, a ray, parallel to the world z axis, is projected from the point toward the collection of triangular surfaces. This line can be expressed as a parametric equation of the form:

$$L = \vec{P_w} + \vec{d_l}t \tag{3.37}$$

Where P_W is a 3×1 vector containing the point's x and y coordinates and the z component is set to zero. d_l is a unit vector denoting the line's direction and is equal to $[001]^T$. The next step is to determine if the line intersects the specific triangle. As is well known from computational geometry [21] a plane Q is uniquely defined by three vertices. This calculation is summarized below for a plane defined by vertices V_1, V_2 and V_3:

$$\vec{N} = \vec{V_1V_2} \times \vec{V_1V_3} \tag{3.38}$$

$$D = N_x V_{1_x} + N_y V_{1_y} + N_z V_{1_z} \tag{3.39}$$

$$Q = N_x x + N_y y + N_z z - D \tag{3.40}$$

At this point, the intersection point of L and Q can be determined by setting equation 3.37 equal to 3.40 and solving for the parameter t[22]:

$$t = \frac{N_x P_{w_x} + N_y P_{w_y} + N_z P_{w_z} + D}{N_x d_{l_x} + N_y d_{l_y} + N_z d_{l_z}} \tag{3.41}$$

Substituting this value for t into equation 3.37 will produce the intersection point for the line and the plane. This calculation assumes that the plane is unbounded, but this is not the case since the vertices V_1, V_2 and V_3 define a triangle. It is therefore possible for the intersection point to lie outside of the triangle. It would be useful to develop a test that can determine if the intersection point lies within the bounds of the triangle. Figure 3.5 depicts a point inside an arbitrary triangle. If the point lies inside the triangle then

$$\angle V_1 P V_2 + \angle V_1 P V_3 + \angle V_2 P V_3 = 2\pi \tag{3.42}$$

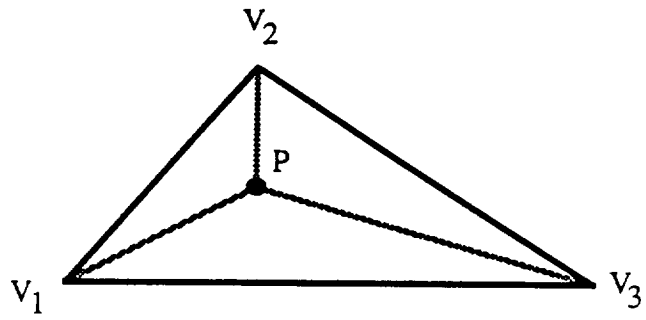


Figure 3.19: Determining if a point lies within a triangle

It is readily apparent that this relationship is only true if P lies inside the triangle or along its edges. Of course, it is assumed that P is coplanar with the triangle.

The method for calculating the intersection point between a line and plane presented here can be applied to any arbitrary line or triangle. This method can be useful in estimating three dimensional points, determining where the laser beam will strike a triangular surface. or if some desired point is occluded by a triangular surface. All this information is possible given that the data points generated by the laser and camera configuration are organized to define triangular surfaces. In effect, by adding structure to the data points, more information can be inferred about a workspace than simply a random collection of points.

3.6 Summary

The techniques presented in this chapter provide a set of useful tools to use a calibrated laser scanner and camera for active visual sensing. As is obvious from the discussions presented in this chapter, these techniques draw from different disciplines. In the case of point estimation, the LSE method is based on statistical mathematics. By contrast, the midpoint to the common normal method as well as

the computer simulation models are premised on geometry. Further, detecting the laser spot draws from image processing methods. Finally, the data structures for organizing three dimensional data is influenced by methods used in computer graphics and computational geometry. The methods presented here provide the means to apply a laser scanner to relevant three dimensional sensing tasks.

4. Application of a Laser Scanner to Three Dimensional Sensing Tasks

The research that has been presented so far describes the individual laser scanner techniques and the experiments used to verify them individually. While it is important to assure that the individual techniques work properly, it is even more important to determine how well these techniques work together to address practical three dimensional visual sensing problems.

Two visual sensing scenarios have been identified that are relevant to the visual sensing requirements of the CIRSSE testbed and can be solved using the laser scanner techniques. The first task uses a laser scanner to calibrate a camera. In this scenario, the well defined properties of the laser beam are used to provide data points to ascertain the camera's intrinsic and extrinsic parameters. The second scenario uses a calibrated laser scanner and camera to generate three dimensional surface maps of the workspace.

This chapter describes how the laser scanner was used to address these visual sensing scenarios. The solutions to the two scenarios draw from the techniques and knowledge presented in the previous two chapters. The application of the techniques will be explained as appropriate. As will become apparent, this chapter emphasizes experimentation and the engineering trade-offs that exist when solving real world problems.

4.1 Calibration of Cameras Using a Laser Scanner

Methods have been developed by Tsai[12], Repko and Sood [15], and Noseworthy[16] to calibrate cameras. These methods determine the camera's intrinsic and extrinsic parameters that best satisfy a collection of data points. A data

C-2

point consists of the coordinates of a three dimensional point and the pixels coordinates of this point when it is projected on the camera's image plane. Repko and Sood produced data points by attaching a target containing several black circles on the end effector of a robot arm. The arm was then moved into the field of view of the camera and the centroids of the circles were identified in the camera image. Noseworthy proposed a variation on this method where a flashlight was placed in the robot's end effector. The end effector was subsequently moved to different locations in the workspace and the centroid of the spot produced by the flashlight was calculated.

These data collection techniques produced reasonably accurate data, but they have several drawbacks. First, since both methods depend on the robot for the coordinates of the three dimensional point, they are both susceptible to kinematic errors. Second, calibrating the cameras is dependent on the availability of the robot. Third, data collection can be slow due to the communication time between the vision and motion control systems and the required time to physically move the robot.

4.1.1 Description of task

To improve on these data collection techniques, a new method was developed which uses the laser scanner to provide the three dimensional points needed for camera calibration. If the laser scanner is calibrated, the laser beam can be directed in a well defined manner. However, directing the laser beam does not, by itself, allow the laser to define a three dimensional point. As was pointed out in section 2.1, one of the axes has to be constrained. If it is assumed that the laser will direct its beam at some plane at a fixed z value, then it is possible for the laser to be directed at any x y coordinate on this plane.

The process of generating data points for camera calibration is depicted in

figure 4.1. A flat planar object is situated at some fixed height above the testbed floor. The laser is then directed at multiple points on this plane. The laser spot reflecting off the plane is detected by the camera, and the centroid of this spot is determined. Since the mirror angles for each point and the z value of the plane are known, the three dimensional coordinates of the laser spot can be determined.

The current implementation of this application directs the laser at sixty-four points on the planar surface, evenly distributed as an eight by eight grid. To calibrate a camera, it is best to collect data points at different z levels in the testbed. This requirement will permit the calibration algorithms developed by Tsai, Sood Repko, and Noseworthy to converge rapidly. Further, if all the data points are coplanar it is possible for the calibration techniques to generate valid camera poses above *and* below the plane of points, which is clearly not correct. To satisfy the requirement, the planar surface is moved to different locations and heights in the testbed. Typically, the planar surface is situated at ten different positions during the calibration process. This results in a total of six hundred forty points.

Some of the techniques developed in chapter 3 can be directly applied to this application. The most obvious technique that is relevant to this task is the direct geometric method for calibrating the laser scanner. Since the laser scanner will be providing the three dimensional point, the laser scanner must be calibrated so the laser beam can be directed accurately. Some of the laser spot detection methods can also be used in this application, particularly spot selections based on size and intensity. While the spot detection algorithms will work under a variety of lighting conditions, optimal performance is achieved when lighting is subdued. This constraint is not unreasonable since the goal of this task is to calibrate the camera and such a task inherently implies that the conditions of this process will be controlled.

The computer simulation model for laser scanner calibration (section 3.3.1) is a valuable tool for camera calibration. The purpose of this simulation is to provide a means to predict the accuracy of the laser scanner in the workspace based on measurement errors incurred during the laser scanner calibration process. The analytical methods used to process the camera data points, will produce more accurate camera calibration results if the three dimensional points are measured accurately. The laser scanner calibration computer model can be used to determine those locations in the testbed where the laser scanner will produce the most accurate results. Hence, the planar surface can be placed at these locations, and as a result, the three dimensional points used to calibrate the cameras will be as accurate as possible.

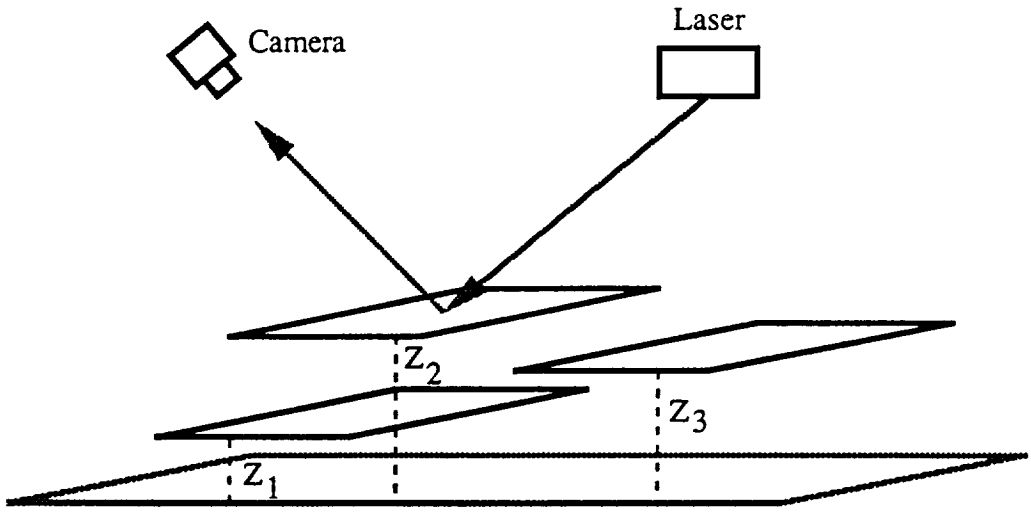


Figure 4.1: Applying a laser scanner to camera calibration

4.1.2 Camera calibration application results

The camera calibration application has been used on several occasions to calibrate the ceiling cameras in the CIRRSE testbed (camera one and camera two). For each calibration, six hundred forty data points were collected for each camera.

Approximately half of the data points was used to calibrate the cameras. The other half was used as control data to verify the calibration. Verifying the calibration was performed by using the camera calibration parameters, the pixel coordinates, and laser scanner mirror angles to estimate the location of the three dimensional points. Estimates were obtained using the laser scanner and camera 1, laser scanner and camera 2, and camera 1 and camera 2. These estimates were then compared to the true locations of the points. Obviously, it is important to determine some means of evaluating the effectiveness of the calibration process. The camera calibrations were evaluated using three criteria:

- *Average 3D Error*: The average three dimensional Euclidean distance between the true position of the three dimensional points and the estimates determined by the laser scanner and cameras. This value provides the expected accuracy of the point estimates determined using the camera and laser scanner.

- *3D Variance*: The square of the standard deviation of the three dimensional Euclidean distance between the true locations of the points and their estimated locations. This value provides a measure of the variability of the accuracy of the point estimates.

- *Max 3D Error*: The largest single three dimensional Euclidean distance between the true locations of a point and its estimated location among all the data points in the test set. This measure provides the worst case value for the point estimates.

The results of two camera calibrations obtained using data points generated by the laser scanner are presented in tables 4.1 and 4.2. There are two comments to make about these results. First, one must be careful in comparing the values in one table with the corresponding values in the other table. The results in each

table are based on different sets of data points and different physical locations of the laser scanner and camera. The results of laser and camera simulations presented in section 3.3.4 clearly indicate that accuracy is a function of the error in the laser and camera parameters and the physical location of the test points in the workspace. Since the calibration results in the tables are based on different sets of data points and different physical arrangements of the devices, the tests results will not be identical. Second, since the purpose of this application is to evaluate the laser scanner techniques under actual conditions in the CIRSSE testbed, the results are based on data collected in the testbed and not on simulated information.

Devices	Avg 3D error	3D Variance	Max 3D error
Laser / Camera 1	4.5mm	12.8mm	19.9mm
Laser / Camera 2	5.7mm	15.7mm	20.6mm
Camera 1 /Camera 2	8.0mm	21.9mm	29.9mm

Table 4.1: Accuracy of camera calibration - June 1991

In general, the average three dimensional error ranged from 5mm to 10mm. Since the distance from the laser scanner and cameras to the test points was approximately 2000mm, these errors constitute 0.25% to 0.5% of the total distance. The three dimensional variance results indicate that there is some variation in accuracy across the data points, but this variation is not severe. The maximum three dimensional error indicated that worst case accuracy constituted between 0.65% and 2.0% of the total distance from laser and camera to the points (assuming that the distance is 2000mm as before).

Devices	Avg 3D error	3D Variance	Max 3D error
Laser / Camera 1	9.7mm	58.2mm	39.8mm
Laser / Camera 2	5.0mm	8.9mm	13.3mm
Camera 1 /Camera 2	5.6mm	9.0mm	16.8mm

Table 4.2: Accuracy of camera calibration - October 1991

Plane	Min X	Min Y	Max X	Max Y	Z
1	1600	-600	2100	-225	0
2	1050	-200	1550	175	230
3	1250	-700	1750	-325	642

Table 4.3: Boundaries of planes of data points used for evaluating the performance of the laser scanner and camera one - October 1991. All units in millimeters.

What was particularly interesting was the results for the laser and camera one presented in table 4.2. These values were higher than all the other values in both tables. It would be instructive to determine why these results occurred. To do this, it is important to examine the data points used for the test. The data set used to evaluate the point estimates generated by the laser scanner and camera one consisted of 192 points distributed over three planar surfaces (sixty-four points per plane). The locations of these planes (in world coordinates) are shown in table 4.3.

During the evaluation of the laser and camera simulations (see section 3.3.4) the camera parameters used for simulations were the same as the camera parameters used to generate the results in table 4.1. For almost all the simulations, the laser and camera experienced maximum error in the rectangular volume bounded in x between 1400mm and 2400mm, in y between -500mm and -1000mm and in z between 0mm and 600mm. Referring to table 4.3, it is obvious that all of plane one and half of plane three fall within in the rectangular region. In short, half of the test points lie in a region where the laser scanner and camera one will exhibit the most error. These results are useful because the laser and camera simulations can be used to interpret performance of the laser and camera under actual field conditions.

While the results of the camera calibration obtained using the laser scanner are reasonably accurate and understood, how do these results compare to those obtained using the data collection methods presented by Sood and Repko and Noseworthy? Sood and Repko reported three dimensional errors between one half and

two millimeters [15]. No values were reported for variance or maximum error. Using data points collected using a robot holding a light source, Noseworthy reported average three dimensional errors between 2.8 and 5 millimeters, variances between 2 and 33mm^2, and maximum three dimensional errors between seven and thirty-three millimeters [16]. Therefore, the camera calibrations obtained using data points generated by the laser scanner are similar to those of Sood, Repko and Noseworthy, but slightly less accurate.

One reason why the calibration results obtained using the laser scanner are slightly worse than using a robot holding a light source or a pattern is due to the estimate of the centroid of the light source or patterns in the camera image. Since the cameras are situated at long distances from the calibration target, it is important to obtain a sub-pixel estimate of the centroids of the light source or pattern. Indeed, at a distance of 2000mm, one pixel in the camera image can constitute four or five millimeters of movement in the workspace. A sub-pixel estimate can be made more accurate if there are more pixels contributing to the estimate. The light source used by Noseworthy and the patterns used by Sood and Repko produce regions in the image plane of one hundred pixels or more. The laser spot generated by the laser scanner consists of approximately thirty pixels. Therefore, there are fewer pixels in the laser spot to contribute to a sub-pixel estimate of the laser spot's centroid.

The camera calibration results obtained using the laser scanner must be kept in perspective. The calibration results obtained earlier were subsequently used in CIRSSE's case study to complete a triangle of struts and nodes. Struts are alluminum rods approximately 60cm long and 2 cm wide while nodes are hexagonal structures that are approximately eight centimeters across, and act as joints to connect struts together. In this scenario the ceiling cameras were used to detect the uncompleted side of the triangle. This information was used to direct the robot to

positions over the uncompleted side so that the cameras mounted on the robot could be used to obtain a more detailed view of the uncompleted portion of the triangle. The point estimates obtained from the ceiling cameras were sufficiently accurate to direct the robot to the correct positions. In short, calibrating the cameras using the laser scanner provided results that were sufficiently accurate to be useful to the research efforts of the CIRSSE testbed.

A key advantage of the laser scanner camera calibration method is it collects data at a faster rate than the other methods used to this point. This speed is due to the fact that the communication time between the laser and camera is shorter than between the camera and the robot. The laser scanner method can collect a data point in approximately 1.5 seconds, while the robot method needs five to ten second per point. Further, it is not dependent on the availability of a robot. Hence, the laser scanner offers a convenient and viable approach to calibrating cameras. The camera calibration results obtained using the laser scanner were also used in a surface mapping application presented in the next section.

4.2 Creating Surface Maps Using a Calibrated Laser and Camera

If the laser scanner and a camera are calibrated to a common world coordinate system, they can be used to generate surface maps. A surface map is defined as a structure that represents a workspace as a three dimensional terrain map. This map is similar to contour maps used in land surveying. A surface map is particularly useful in that it can provide information to safely guide a robot through a workspace. The surface map application presented here draws from the methods and techniques developed in chapters 2 and 3. This section provides a detailed description of the surface mapping task and the problems encountered in its development. Additionally, results of the surface mapping application will be presented.

4.2.1 Development of surface map application

A laser scanner and camera can generate surface maps by directing the laser beam into the workspace and using the camera to detect the laser spot. This process is repeated until a collection of three dimensional points are obtained. The range points are then stored to facilitate analysis and enhancement. This brief description of the task is simplistic in that it does not highlight the engineering problems. In order to develop the application, several issues have to be addressed:

1. How will the laser beam be directed through the workspace?
2. How will the range estimates from the laser and camera be determined?
3. How will the three dimensional surface map be constructed from the data points?
4. How can the surface map be analyzed, enhanced or displayed?

Developing a method to direct the laser beam through the workspace is necessary to ensure that the surface map is generated efficiently. To control the laser scanner a triangular plane is defined that encompasses a portion of the workspace. This plane is parallel to the xy plane of the world coordinate system. The triangular plane is recursively divided into smaller triangles using the techniques describe in section 3.5. This produces a list of *target points* and a quad tree that describes how these points form triangular surfaces. The laser is then directed at each point in the list. It should be emphasized that these target points are not intended as range data. In practice, when the laser beam is directed toward a target point, it will strike an object before or after the beam passes through the point. The range data (or *range point*) will be the location of the laser spot when the beam strikes an object. Hence, the target point serves as a convenient means of defining a region of the workspace to be scanned by the laser.

The next issue in developing the surface map application is how to obtain range estimates using the camera and laser scanner. This problem consists of two tasks: 1) detecting the laser spot 2) estimating its position. In the previous chapter, a great deal of research was conducted to develop methods for detecting a laser spot in the camera image. Since the surface map application should operate under the widest variety of conditions, all four criterion tests are employed for detecting the laser spot. Because the different tests possess different computation times, the tests are executed in the following order:

1. elimination by intensity
2. elimination by size
3. elimination by movement
4. elimination by triangulation

The order of these tests is designed to maximize the efficiency of the laser spot detection process. As mentioned in section 3.4.5.4, the size and intensity tests exhibit the smallest computation; so they are executed first to reduce the size of the region list. The movement test does require comparing region data between successive frames, but since the surface map application is collecting a set of data points, two region lists can be stored easily and a simple software switch can alternately load new region data into either list. Since the computation time of the movement test is greater than either size or intensity, but less than the triangulation test, it is the third test executed. The last test to be executed is the triangulation test.

Another problem in determining range estimates is selecting the point estimation algorithm that will generate the range coordinates. The simulations presented in section 3.3.3 indicate that the LSE and midpoint to the common normal methods

yield comparable results. It would seem, therefore, that either method would be suitable. However, the LSE point estimation method exhibits better performance if the rotation matrix of the camera's extrinsic parameters is not orthonormal. This phenomenon was discussed previously in section 3.2 with regards to point estimation methods. When the surface mapping application was under development, there were two methods for calibrating the cameras. One method is based on a calibration method developed by Tsai[12] and does not guarantee that the rotation matrix will be orthonormal. The other method, developed by Noseworthy[16], guarantees that the rotation matrix will be orthonormal. Both methods are currently in use at CIRSSE. Since the LSE method can better accommodate distortions in the rotation matrix, it was decided that the LSE method should be used for calculating range estimates. In this way, the surface mapping application can accommodate camera parameters generated by either camera calibration method.

The third issue in developing this application is coming up with a method for constructing a surface map based on the range data collected using the laser scanner and the camera. Solving this problem is straightforward. As mentioned previously, when the list of target points is generated, the quad tree relating these target points is also created. When the laser beam is directed at each target point, the coordinates of the resulting laser spot is determined, and these coordinates are substituted for the corresponding target point. As a result, the quad tree and the range data can be combined to generate the surface map.

The final issue in developing the surface map application is developing methods for analyzing and enhancing three dimensional data. Since the surface map consists of a collection of triangular surfaces, the most obvious approach is to analyze the surface map using techniques used in computer graphics and finite element meshs. This approach, however, is a bit constraining since it limits the available techniques.

If the structure can be transformed then it would be possible to use a greater variety of methods to process the results. Indeed, if the range data can be transformed into something resembling a two dimensional intensity image, standard image processing techniques can be used on the range data. A method was developed that transforms the surface map into a matrix of height values. The process is performed as follows:

1. Define a matrix of points in a plane that is parallel to the world's xy plane. The points in this matrix should be evenly spaced along the x and y axes.

2. For each point in the plane, project a line perpendicular to the plane and passing through the point.

3. For each line projected from the plane, traverse the quad tree to identify the smallest triangular plane that intersects the line segment. Enter the height of this intersection point into the matrix. The test used to determine if a line and a triangle intersect is described in section 3.5.

4. Save the matrix of points to a data file. Since the points in the matrix are evenly spaced and the spacing is known, it is possible to determine the three dimensional coordinates of each point in the matrix.

This procedure indicates that the quad tree is searched until the smallest triangle that intersects the line is line is found. While this will provide the most accurate estimate of the height of the map at the given intersection point, It is not always necessary to have such accuracy. Each level of the quad tree represents one level of division of the original triangular plane, and the smallest triangles are located at the leaf nodes of the tree. The procedure has been modified to either search the quad tree to its full height, or stop at some predefined level and use the height of the intersection of the line with the triangle at that level. The procedure for

converting data points also provides a means to display the surface map. Since the data is converted to a matrix form, mathematical analysis software such as Matlab and DeltaGraph can be used to render the data as a three dimensional mesh.

It is obvious from this discussion, that the surface map application developed for this research draws heavily from the techniques and knowledge obtained from the previous chapter. To clarify the exact structure of the surface map application resulting from this discussion, the steps used in creating the surface map are enumerated below:

1. Define a triangular plane somewhere in the workspace. Generate a list of target points and a quad tree of triangles.

2. Use the list of target points to direct the laser beam. Estimate the three dimensional coordinates of the resulting laser spot for each target point. Replace the target points with the range data.

3. Using the quad tree and range data, create a surface map of triangular surfaces. Use this surface map to generate a matrix of height values.

4.2.2 Results of surface map application

The surface map application was used to map different objects. The first test was to scan the laser over a table top situated at approximately 635 millimeters above the testbed floor. While the resulting surface map is simply a flat plane, the goal of this test was to determine the variability of the range data over the entire table top. The results of this test are summarized in table 4.4. The range data exhibited a mean z value of 623mm which is 12mm below the true height of the table. Since the laser and camera are approximately 2000mm from the table top the observed error constitutes 0.6% of the total distance. What is more interesting is that the

Total points	1026	Table height	635mm
Max height	632mm	Min Height	611mm
Mean height	623mm	Std. Deviation	3.75mm

Table 4.4: Results of range data on flat fixed plane

standard deviation was 3.75mm which means that the error is relatively constant over the entire table top. It is this second fact which is important since inconsistent performance would greatly affect the reliability of the surface map application.

During the second test, a stack of books and a cylindrical object were placed on the table. These objects are presented shown in color plate A.3 in appendix A. The books were stacked in a circular arc and staggered so that part of each book in the stack was visible to the laser scanner and camera. The surface map application was configured so that the minimum edge length of the triangles in the surface map was 15mm. A matrix of height values was generated using the surface map. The spacing between points in the matrix was set to 6mm in both the x and y directions. The unfiltered matrix of height values for this set of objects is depicted in figure 4.2. The individual books in the stack are clearly visible as is the circular arc shape of the stack. The cylinder is recognizable although the top is a bit uneven. It is also apparent that there are some anomalies in the surface map. Approximately 3000 points were used to generate the map. Of these points, two or three hundred generated no laser spot and were marked as not found. As a result the surface map does exhibit some anomalies. To reduce the effect of these anomalies, the matrix was convolved with a 3×3 smoothing kernel. The results of this filtering are shown in figure 4.3. The anomalies are greatly reduced, but the filter (which is effectively low pass) has, as expected, reduced the shear slope of the cylinder walls and the side of the stack of books. However the reduction in slope of these features is not severe.

The surface maps presented in figures 4.2 and 4.3 were generated by traversing

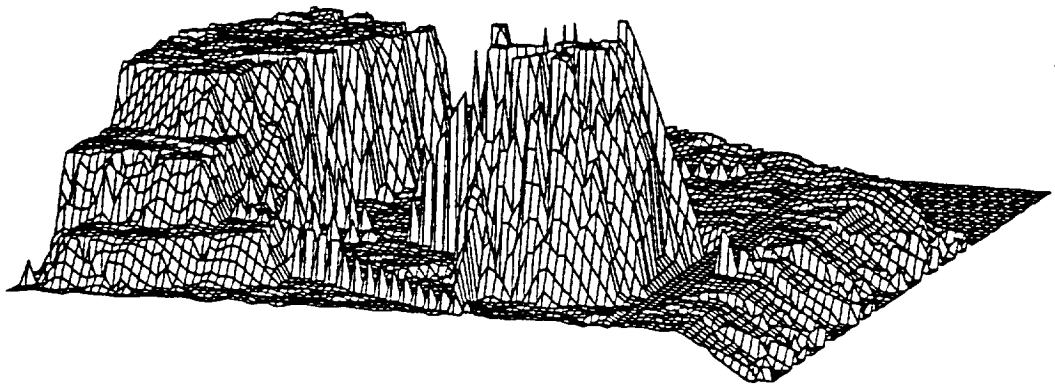


Figure 4.2: Unfiltered surface map of a stack of books and a cylindrical object

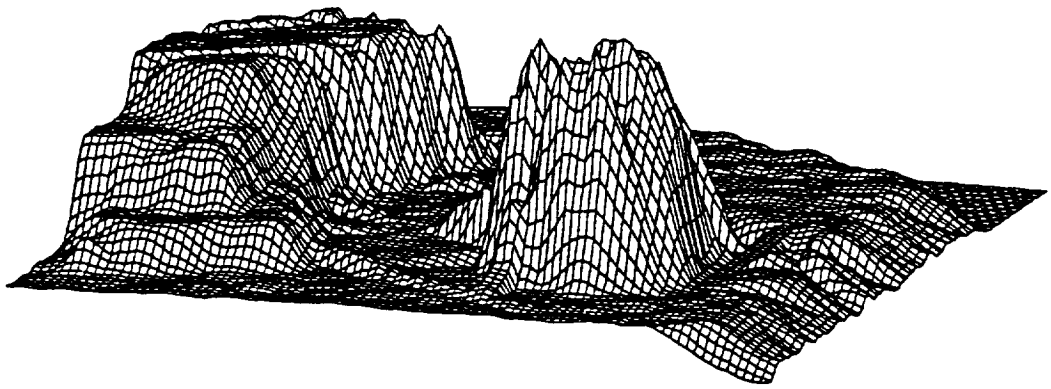


Figure 4.3: Surface map convolved with 3×3 low pass filter

the full depth of the quad tree of triangles (there were eight levels in the quad tree for this example). As mentioned previously, it is possible to generate a surface map by traversing the quad tree to a specific depth. Figures 4.4 - 4.7 show the surface maps that result from traversing the quad tree to a depth of one level (the root node), three levels, five levels and seven levels. As the quad tree is traversed to higher levels, the amount of detail in the books and cylinder increases. Indeed, by the fifth level, the major features of the books and cylinder are readily visible.

The results of the surface map application indicate that the laser scanner can be a useful tool for three dimensional visual sensing. The current implementation highlights some of the potential problems associated with generating surface maps with the laser scanner. First, if the laser strikes an objects at an oblique angle, the laser spot can become severely distorted. This distortion takes the form of a long irregular streak that does not conform to any consistent geometry. In such cases, the camera is unable to detect the laser spot, and since the distortion is so irregular and unpredictable, it is difficult to compensate for. Another exceptional case arises when the laser strikes a highly reflective object. In such cases, the reflection generates a laser spot and one or more secondary reflections. In these cases, more than one laser spot is detected, and since it is impossible to obtain a unique solution, the point is marked as not found.

Despite these irregularities, the surface maps that are generated by the laser scanner and camera are of sufficient fidelity to be useful for three dimensional sensing. It is particularly useful when objects in a scene are of similar intensity. A purely passive camera approach might no be able to distinguish the boundaries of the objects, but a laser scanner and camera configuration is not so readily deceived since the laser scanner can generate a laser spot that the camera can detect. Additionally, all of the surface maps presented in this section were generated under

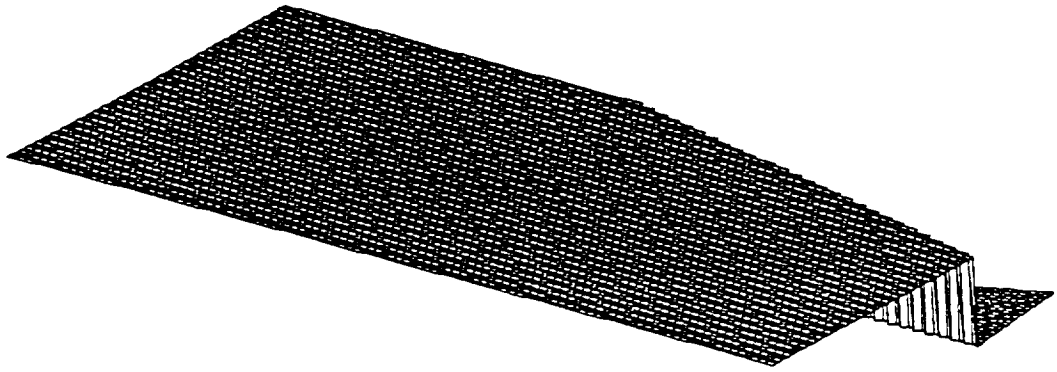


Figure 4.4: Surface map obtained by traversing quad tree by one level

normal lighting conditions in the CIRSSE testbed. In other words, no special lighting arrangements were made to conduct these experiments. This fact speaks to the flexibility of the laser spot detection algorithms discussed in section 3.4. Indeed, a special case of this range map experiment was conducted in which odd pieces of unfinished metal were placed into the scene. This scenario is shown in color plate A.5. The camera image of this scene is shown in color plate A.6. The specular reflections from the metal objects fall within the same size and intensity criteria as the laser spot, and hence, have been marked with blue crosses. However, the movement and triangulation tests are able to screen out the extraneous regions and correctly acquire the laser spot (indicated with a red cross).

4.3 Summary

This chapter has presented two situations where the laser scanner techniques could be applied to solve visual sensing problems. The first scenario was utilizing a laser scanner to calibrate cameras. The results of this application indicate that

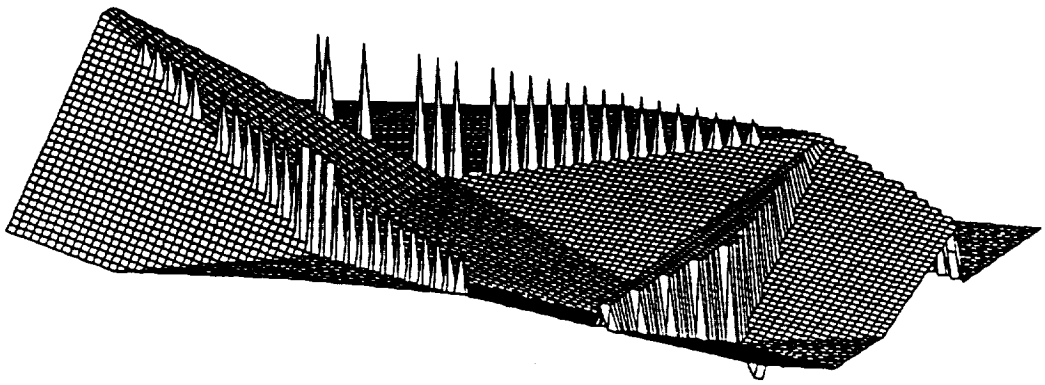


Figure 4.5: Surface map obtained by traversing quad tree by three levels

the laser scanner can collect data more rapidly than alternate methods that utilize a robot arm, yet produces camera calibrations that are sufficiently accurate to be useful in CIRSSE's research efforts. The laser scanner calibration techniques were necessary to permit the laser scanner to generate three dimensional points. The laser scanner calibration simulation provided a means to determine where in the workspace the laser scanner would produce the most accurate points. The size and intensity tests used for detecting the laser spot were necessary for data collection, and the laser and camera simulation provided a means to understand and evaluate the accuracy of the final camera calibrations.

The second scenario utilized a laser scanner and a camera to generate surface maps of objects in the workspace. The results of this application provide a means to create a three dimensional map of the workspace based on a set of range points. This application used the laser scanner calibration techniques developed in chapter 2. All the laser spot detection methods (size, intensity, movement, and triangulation)

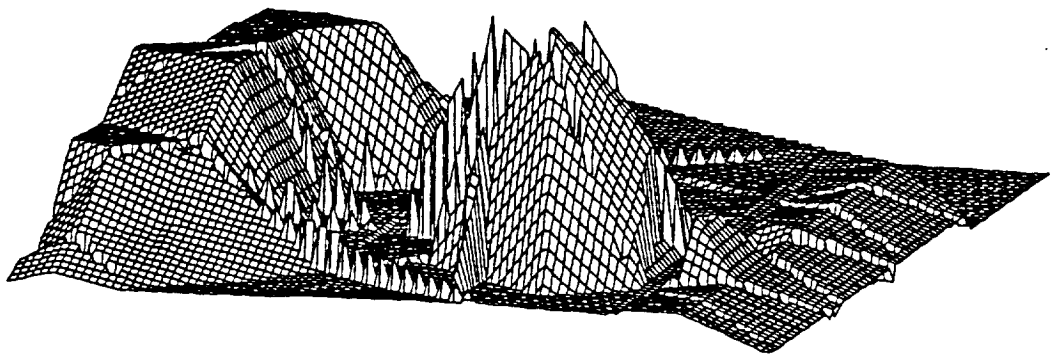


Figure 4.6: Surface map obtained by traversing quad tree by five levels

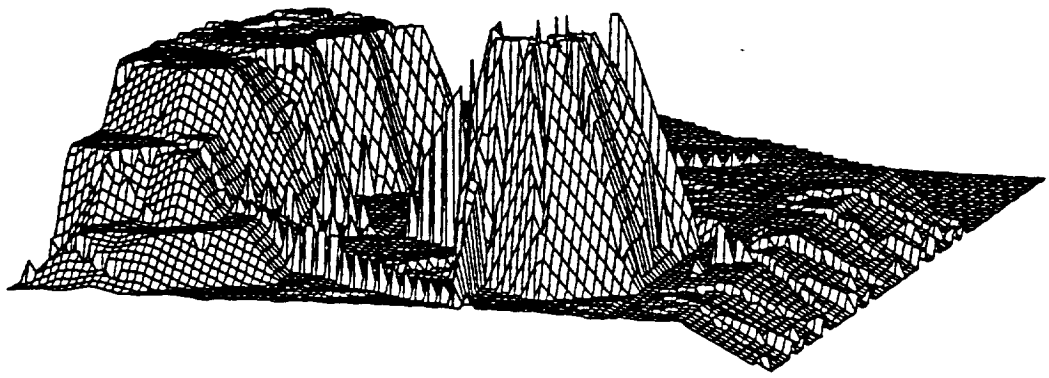


Figure 4.7: Surface map obtained by traversing quad tree by seven levels

developed and evaluated in section 3.4 were used to permit the laser spot to be detected under normal lighting conditions. Finally, the triangular representation of data points discussed in section 3.5 provided three dimensional structure to range data obtained using the laser scanner and camera.

In closing, the applications presented here indicate that the laser scanner techniques can be used together to solve relevant three dimensional visual sensing problems. This is not to say that the laser scanner is the only form of three dimensional sensing. Indeed, it is important to evaluate the utility of the laser scanner in the context of other methods of three dimensional sensing.

5. Appraisal of Laser Scanner for Three Dimensional Sensing

The applications of a laser scanner presented in the previous chapter illustrate how the techniques developed for this research can be applied to solve visual sensing problems. At this point, it is appropriate to appraise the performance of the laser scanner techniques developed for this research in the context of alternate visual sensing methodologies. In performing this appraisal, the laser scanner techniques will first be evaluated with respect to other active sensing methods. Subsequently, the laser scanner techniques will be contrasted with common passive sensing methods. Finally, the laser scanner techniques will be evaluated in the context of a visual sensing system that employs several different visual sensing methodologies.

5.1 Laser Scanning in the Context of Active Visual Sensing

Active visual sensing is characterized by the generation of some well defined light signature (structured light) and directing this light into a workspace. A camera is then used to detect the reflection of the structured light as it strikes objects in the workspace. The differences among the active sensing techniques pertain to the form of the structured light generated and the image processing algorithms used to recover three dimensional data from the camera image. Most active sensing techniques can be grouped into three major categories. The first group includes methods that generate a single beam of light. The second group of methods generate a "sheet" or plane of light, and the third group generates some predefined pattern of light such as a grid.

Active sensing methods that generate a single light beam operate by first directing the beam of light into the workspace. The reflection of this beam as it strikes an object is then detected by a camera and stereoscopic triangulation can be

employed to determine the three dimensional location of the reflection. The laser scanner techniques presented in this report fall under this category of active visual sensing.

There are several advantages to this mode of visual sensing. First, the reflection of the light beam is a simple signature to detect as opposed to a complex pattern. Second, the light source is focused into a single beam of light, and, hence, is delivered efficiently into the workspace. This efficiency is increased if the light source is a laser since there is less beam dispersion with a laser than with a light source such as a spotlight [23]. Efficiency is important because the reflection of the beam can be detected at longer ranges if more photons contribute to generating the reflection[24]. Additionally, increased efficiency implies less illuminating power to accomplish the same task and such savings can result in lighter, smaller components or lower demand on the electrical facilities powering the vision system. Such savings are critical on space based platforms since size, weight and power are important concerns.

This mode of visual sensing is not without problems. For every camera image acquired, only one three dimensional point can be obtained. It is for this reason that three dimensional structuring, such as triangular planes (see section 3.5), are useful in maximizing the information obtained from each camera image. However, a more complex structured light pattern can potentially yield more three dimensional data from each camera image.

A second active sensing approach involves projecting a plane (or sheet) of light into the workspace. This plane of light will produce line segments (or stripes) when it strikes objects in the workspace. A camera is again used to observe the workspace and identify the line segments. There are several ways to generate a plane of light. The most common method is to use a projector to generate a slit of

light that is subsequently projected onto a mirror which directs the sheet of light into the workspace[25]. Alternatively, a sheet of light can be produced by passing a laser beam through a cylindrical lens[2].

Since a line of structured light appears in the image plane instead of a single point, more three dimensional information can be obtained for each camera image. However, the light source distributes its power over a line segment which means that the effective operating distance of this visual sensing configuration is less than if the light was concentrated into a single beam. Another difficulty with using a sheet of light is detecting the line projected on objects in the workspace. The line can be broken if it projects on multiple objects, or it can be curved if the objects have curved surfaces. In order to effectively identify the structured light, either the camera is fitted with a bandpass filter that is sensitive to the structured light (a simple matter is a laser is used as the light source), or the ambient light is controlled so that the structured light creates the brightest regions in the image[26].

Another active visual sensing method projects a pattern of light into the workspace. This pattern is typically a grid [27] or multiple stripes. A camera detects this pattern and measures the pattern's distortion to obtain three dimensional information. The grid is most frequently generated using a projector and a slide containing the pattern.

The pattern can be dispersed over the entire workspace or a small pattern can be projected and subsequently moved through the workspace using a panning mirror. This method has one advantage in that projecting a pattern implies that more three dimensional information can be extracted from each camera image than is possible with a single beam of light or a stripe. Of course, the computational complexity increases, but this cost should be offset by not having to acquire multiple images.

There are several problems with this approach. First, the pattern is more

complex than a single point or sheet of light and the task of detecting the distorted pattern maybe extremely difficult in some cases. This problem is compounded if the grid is incomplete due to occlusion. Second, the illumination power generated by the light source is distributed over the entire patterns. Hence, the intensity of the pattern in the camera image will be reduced, thereby decreasing the effective operating distance of the vision system. As with the case with the sheet of light, most pattern methods either use a camera fitted with a bandpass filter or ensure that the pattern is the brightest light in the workspace by reducing ambient light.

With these different alternatives, there is the inevitable question of which method is best. The answer to this question depends on the environment in which the vision system will operate. In a laboratory setting, or an industrial assembly line, a sheet of light or a pattern is useful since lighting conditions can be controlled and it is important to maximize the amount of three dimensional data obtainable in one camera image. In a space environment, however, lighting conditions are harder to control and the distances between the vision system and objects in the scene will be longer than in a laboratory. Hence, a single beam of light may be more useful since its concentrated beam can be detected over longer ranges than a sheet of light or a pattern.

5.2 Laser Scanning in the Context of Passive Visual Sensing

Passive visual sensing relies on extracting three dimensional information based entirely on the information available in the scene. Passive sensing methods can be divided into two major groups. The methods in the first group use fiducial marks on objects in the scene to aid in identification. A second method identifies features in the scene using multiple cameras and attempts to relate these features to some model representation available to the vision system.

The first passive visual sensing method employs fiducial markings on objects in the workspace. Dual cameras are then used to identify the fiducial marks and stereoscopic triangulation is used to determine the three dimensional coordinates of the object. There are no constraints on the form of a fiducial mark. Fiducial marks can simply identify an object or they can be designed to provide position and orientation information about the object. Further, a fiducial mark can be a passive feature such as a pattern imprinted onto an object, or it can be a light that acts as a beacon similar to running lights on a ship or airplane.

No matter what method is used, the premise of the fiducial mark approach is to identify an object by placing some of the burden of identification onto the object and not the vision system. Using fiducial marks has several advantages. The most important advantage is that the vision system can be optimized to detect the fiducial marks. While this dependence on the fiducial marks has advantages, it does pose several problems. First, if an object bearing a fiducial mark is far away from the vision system or oriented in some position, the fiducial mark may be unreadable. Second, adverse lighting conditions may prevent the vision system from identifying the fiducial mark. Third, if the fiducial mark is damaged, the vision system may not be able to recognize it. In a space environment, it is possible that a fiducial mark could be damaged.

Far and away the most prevalent configuration for three dimensional visual sensing is two or more cameras in a completely passive mode. The images produced by each camera are processed to identify specific regions or features. The features located in the individual camera regions are then correlated and stereoscopic triangulation is employed to determine the three dimensional structure of the objects corresponding to the features. A database of three dimensional objects is then consulted to identify the objects. This database can be hard coded to a specific task,

or be more general purpose.

This method can provide the most three dimensional information of all the methods discussed. This method, ideally, does not rely on fiducial marks. In fact, the only presumption made about the object in the scene is that it is described in the database.

Ideally, this visual sensing method has so much potential that it seems to be the best sensing configuration for most situations. While this is true for the ideal system, in reality there are some problems. For example, adverse lighting conditions can severely hamper the performance of the vision system. In the worse case, there may be no light at all in which case the vision system is effectively blind. Further, occlusion of objects by other objects can make identification difficult. Additionally, how the objects should be represented in the model database is not entirely clear. In short there are several problems that must be overcome in order to realize the full potential of this approach and solutions to these problem are not trivial.

In light of the passive sensing techniques just described it is instructive to contrast the merits and limitations of the laser scanner sensing techniques. If lighting conditions are such that the passive multiple camera system can perform properly, then there is no doubt that the multiple camera approach is far and away superior to the laser scanner. However, if lighting conditions are poor, then the laser scanner becomes a better technique. The laser scanner technique is, in a way, similar to the fiducial mark approach in that the reflection of the laser beam is a fiducial mark except that it is generated by the vision system and not physically attached to the object.

5.3 Laser Scanner as a Supplement to an Integrated Vision System

The discussion up to this point has consisted of an adversarial comparison between a laser scanner and other active and passive visual sensing techniques. Such a comparison is important to provide a means of identifying the strengths and limitations of each method. However, a vision system that is capable of operating reliably under a variety of environmental conditions should incorporate more than one visual sensing technique.

Given the sensing configurations mentioned in the previous sections, which methods should be combined to create a vision system suitable for the space environment? Because of its potential, the multiple camera system that relates objects to a model database should form the cornerstone of the vision system. As mentioned previously, relating objects to a database can be difficult due to adverse lighting or occlusion of objects. In a space environment, most of the objects that will be encountered will be man-made. Hence, it is not unreasonable to *a priori* place fiducial marks on some of these objects to aid in identification and determination of pose.

The use of fiducial marks would assist in alleviating some of the object identification problems for the vision system, but it does not fully address the issue of adverse lighting. An active sensing capability will be necessary to illuminate targets when lighting conditions degrade the vision system's ability to identify objects. Of the active sensing techniques presented previously, the single beam of light approach has the most potential. This choice is based on several reasons. First, assuming all things being equal, the effective range for this configuration is greater for a single beam of light than for a pattern or sheet of light. Second, the reflection of a single beam of light off of an object is easier to detect than a strip or a pattern. These two characteristics are important particularly in poor lighting conditions. If lighting conditions are such that the vision system cannot identify objects reliably, it is

important to have the capability to inject ground truth into the scene that is readily detectable and effective over a large area of the scene.

In appraising the utility of the techniques developed under this research effort it is clear from the preceding discussion that an active sensing configuration incorporating a laser scanner can serve as part of a larger vision system that incorporates several visual sensing techniques. In this context, a laser scanner can be a valuable tool for visual sensing.

6. Conclusions and Future Research

In order to draw conclusions about the research presented in this paper it is appropriate to examine the work performed for this research in the context of the objectives laid out in section 1.2. These objectives are again summarized here along with a brief description of the research that was performed to fulfill each objective:

1. Identify the mathematical relationships that govern the operation of a laser scanner.

   This objective was accomplished in section 2.1 where the relationship between a three dimensional point defined in the laser's coordinate space and the orientations of the scanning mirrors was described. The relationship that was presented accounts for the pincushion distortion inherent in a laser scanner.

2. Calibrate the laser scanner.

   Two methods were presented to calibrate a laser scanner. The first method uses an analytical approach that expresses the calibration process as an overdetermined set of linear equations. The second method takes advantage of the inherent geometry of the calibration problem and measures the Euler angles of the laser scanner and its translational position with respect to a known world coordinate system.

3. Develop methods to estimate three dimensional points using a calibrated laser scanner and a camera.

   Two methods were presented for addressing this problem. The first method calculates a least squared error solution that best satisfies an overdetermined system of linear equations that describe the relationship between the three

dimensional point, pixel coordinates of the projection of the point in the camera's image plane, and the scanning mirror angles required to direct the laser beam at the point. A second method projects rays from the laser scanner and the camera into space and calculates the midpoint to the common normal of these rays. Both methods were found to yield comparable results in most cases.

4. Develop computer simulation models of the laser scanner and the laser / camera sensing configurations.

   A laser scanner simulation was presented in section 3.3 that can predict the accuracy of the laser scanner calibration based on estimates of measurement errors during the calibration process. This simulation uses the direct geometric methods for laser scanner calibration. The laser / camera simulation provides a means to estimate errors in system performance when either the camera or laser scanner are rotationally or translationally perturbed.

5. Develop methods to identify the laser spot in a noisy camera image.

   A heuristic method for identifying a laser spot was presented in section 3.4 as well as the results of a series of tests to estimate the performance of this method. Accomplishing this objective is particularly significant since it provides the means to overcome the restriction that the laser spot be the brightest region in the camera image [9]. By overcoming this restriction, the laser scanner scan be used in a wider variety of lighting conditions and scene configurations.

6. Identify data structures and models for representing three dimensional information.

The techniques used in this research to address this problem draw from the disciplines of finite element meshs and computer graphics. The range data acquired using the laser scanner and camera is organized into triangular planes. The planes form a hierarchy that can be represented as a quad tree. This type of data structure provides the means to infer more three dimensional data than would normally be possible with a simple collection of points.

In addition, two three dimensional sensing applications were developed to evaluate the techniques developed to accomplish the research objectives. The first application uses a calibrated laser scanner to calibrate cameras. The second application uses a laser scanner and a camera to generate surface maps. The results of these applications indicate that the individual techniques developed for this research operate properly and they can be combined to accomplish useful visual sensing tasks.

Finally, the laser scanner was evaluated with respect to other active and passive visual sensing techniques. It is clear that a laser scanner cannot, by itself, solve the general visual sensing problem. However, it can be a useful part of a system that employs several visual sensing techniques, and it can do this by providing an active sensing capability that operates over a large work volume and whose signature is readily detectable.

Overall, the objectives of this research have been successfully accomplished. As a result, the CIRSSE testbed possesses an active visual sensing capability that is well understood and relevant to CIRSSE's visual sensing requirements. As with most research, the laser scanner work presented here highlighted several opportunities for future research.

One avenue of future research is a closer examination of the effects of lighting on visual sensing systems. This will be an important factor for any vision system

that is ultimately used in a space environment since lighting conditions will change frequently and often unpredictably.

Another potential area of research is to enhance the utility of the laser scanner by fitting a robot arm with a laser scanner. Small penlight sized lasers are currently in production and it may be possible to use such a laser in conjunction with a scanning mechanism to provide a laser scanner at the end effector of a robot. When used in conjunction with the cameras already mounted on the robot arm, this laser scanner could be used to detect small details of the scene or direct structured light into areas that are not accessible to the laser scanner situated above the testbed. The techniques developed for this report could readily be adapted to a robot mounted laser scanner. However, there are two primary technical problems that would have to be addressed to construct a laser scanner on the robot arm. First, the system must be compact and rugged. Second, the scanning mechanism must be capable of dealing with vibrations as the robot arm is moved.

In closing, this research provides CIRSSE with an active sensing capability. Addressing the engineering problems associated with this active sensing capability has led to a deeper understanding of its capabilities and limitations. Finally, this research has revealed additional opportunities for vision research that are relevant to robotics in space applications.

7. LITERATURE CITED

[1] A. B. Colquhoun, D. W. Cowan, and J. Shepard, "Trade-offs in rotary mirror scanner design," in *Beam Deflection and Scanning Technologies, SPIE Proceedings*, (San Jose, California), p. 12, February 1991.

[2] R. J. Ahlers and J. Lu, "Stereoscopic vision - an application oriented overview," in *Optics, Illumination, and Image Sensing for Machine Vision IV, SPIE Proceedings*, (Philadelphia, Pennsylvania), pp. 298 – 308, November 1989.

[3] G. F. Marshall, T. J. Vettese, and J. H. Carosella, "Butterfly line scaner: Rotary twin reflective deflector that desensitizes scan-jitter to wobble of rotaional axis," in *Beam Deflection and Scanning Technologies, SPIE Proceedings*, (San Jose, California), p. 37, February 1991.

[4] C. J. Kramer, "Holographic deflectors for graphic arts applications: an overview," in *Beam Deflection and Scanning Technologies, SPIE Proceedings*, (San Jose, California), p. 68, February 1991.

[5] H. DeWeerd, "Compact, low power precision beam steering mirror," in *Beam Deflection and Scanning Technologies, SPIE Proceedings*, (San Jose, California), p. 207, February 1991.

[6] G. C. Loney, "Design and performance of a small two-axis high-bandwidth steering mirror," in *Beam Deflection and Scanning Technologies, SPIE Proceedings*, (San Jose, California), p. 198, February 1991.

[7] K. Pelsue, "Precision, post-objective, two-axis, galvanometer scanning," in *High Speed READ/Write Techniques for Advanced Printing and Data Handling, SPIE Proceedings*, (Los Angeles, California), p. 70, January 1983.

[8] J. S. Erhrmann, "Optics for vector scanning," in *Beam Deflection and Scanning Technologies, SPIE Proceedings*, (San Jose, California), p. 245, February 1991.

[9] J. R. Noseworthy and L. A. Gerhardt, "Three dimensional vision - requirements and applications in a space environment," in *Symposium on Advances in Intelligent Systems, SPIE Proceedings*, (Boston, Massachusetts), p. 1387, November 1990.

[10] J. Watson, "Testbed kinematics and routines." CIRSSE Technical Memorandum, March 1991.

[11] W. J. Mistretta, "Analysis, enhancement, and application of a three dimensional inspection system in computer integrated manufacturing," Master's thesis, Rensselaer Polytechnic Institute, Electrical & Computer Systems Engineering Dept., Troy, NY 12180-3590, December 1989.

[12] R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," *IEEE Journal of Robotics and Automation*, vol. RA-3, pp. 323–344, August 1987.

[13] J. J. Craig, *Introduction to Robotics, Mechanics, & Control.* Reading, MA: Addison-Wesley Publishing Company, 1986.

[14] W. H. Press, B. R. Flannery, S. A. Taukolsky, and W. T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing.* Cambridge University Press, 1990.

[15] D. Sood, M. C. Repko, and R. B. Kelley, "An implementation of a camera technique to obtain three-dimensional (3D) vision information for simple assembly tasks," CIRSSE Report 76, Rensselaer Polytechnic Institute, Troy, NY, November 1990.

[16] J. R. Noseworthy, "Inaccuracies of three dimensional vision systems - theory and practice," Master's thesis, Rensselaer Polytechnic Institute, Electrical & Computer Systems Engineering Dept., Troy, NY 12180-3590, August 1991.

[17] R. J. Schalkoff, *Digital Image Processing and Computer Vision.* New York, NY: John Wiley & Sons, Inc., 1989.

[18] D. H. Ballard and C. M. Brown, *Computer Vision.* Englewood Cliffs, NJ: Prentice-Hall Inc., 1982.

[19] W. J. Schroeder and M. S. Shephard, "On rigorous conditions for automatically generating finite element meshes," in *Product Modelling for Computer-Aided Design and Manufacturing* (J. Turner, J. Pegna, and M. Wozny, eds.), Elsevier Science Publishers B. V., 1991.

[20] M. S. Shephard, "Approaches to the automatic generation and control of finite element meshes," *Applied Mechanics Reviews*, vol. 41, pp. 169 – 184, April 1988.

[21] M. L. Boas, *Mathematical Methods in the Physical Sciences.* John Wiley & Sons, 1983.

[22] J. D. Foley, A. Van Dam, S. K. Feiner, and J. F. Hughes, *Computer Graphics - Principles and Practice.* Addison-Wesley Publishing Company, 1990.

[23] Y. F. Wang and J. K. Aggarwal, "An overview of geometric modelling using active sensing," *IEEE Control Systems*, vol. 8, pp. 5 – 13, June 1988.

[24] J. A. Jalkio, R. C. Kim, and S. K. Case, "Three dimensional inspection using multistripe structured light," *Optical Engineering*, vol. 24, pp. 966–974, November 1985.

[25] K. Boyer and A. Kak, "Color-encoded structured light for rapid active ranging," *IEEE Transactions on Pattern Recognition and Machine Intelligence*, vol. PAMI-9, pp. 14 – 28, January 1987.

[26] P. Saint-Marc, J. L. Jezouin, and G. Medioni, "A versatile PC-based range finding system," *IEEE Transactions on Robotics and Automation*, vol. 7, pp. 250 – 256, April 1991.

[27] G. C. Stockman, S. W. Chen, G. Hu, and N. Shrikhande, "Sensing and recognition of rigid objects using structured light," *IEEE Control Systems*, vol. 8, pp. 5 – 13, June 1988.

APPENDIX A
COLOR PLATES

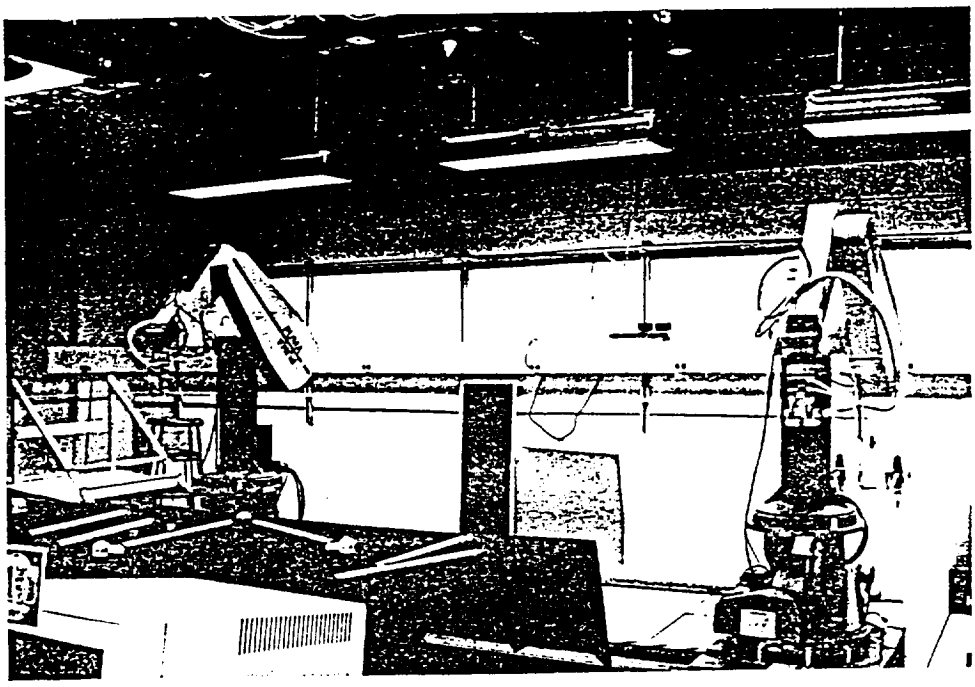


Figure A.1: CIRSSE robot testbed. Struts and nodes are located on the black table. Note red spot generated by the laser scanner reflecting off of the node on the table in the lower left quadrant of picture

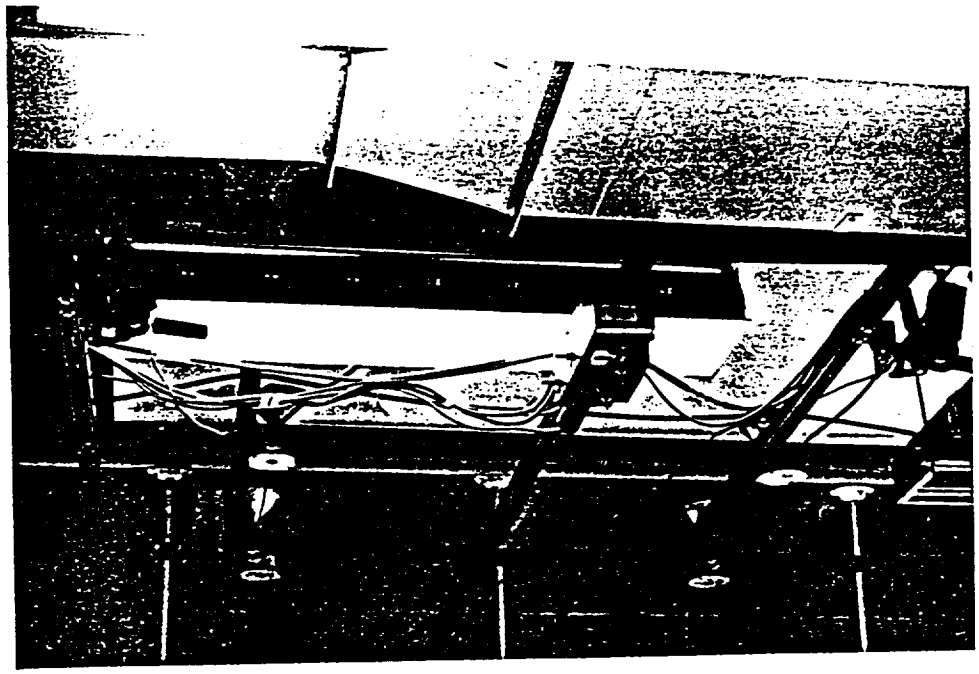


Figure A.2: Ceiling cameras and laser scanner situated over CIRSSE testbed

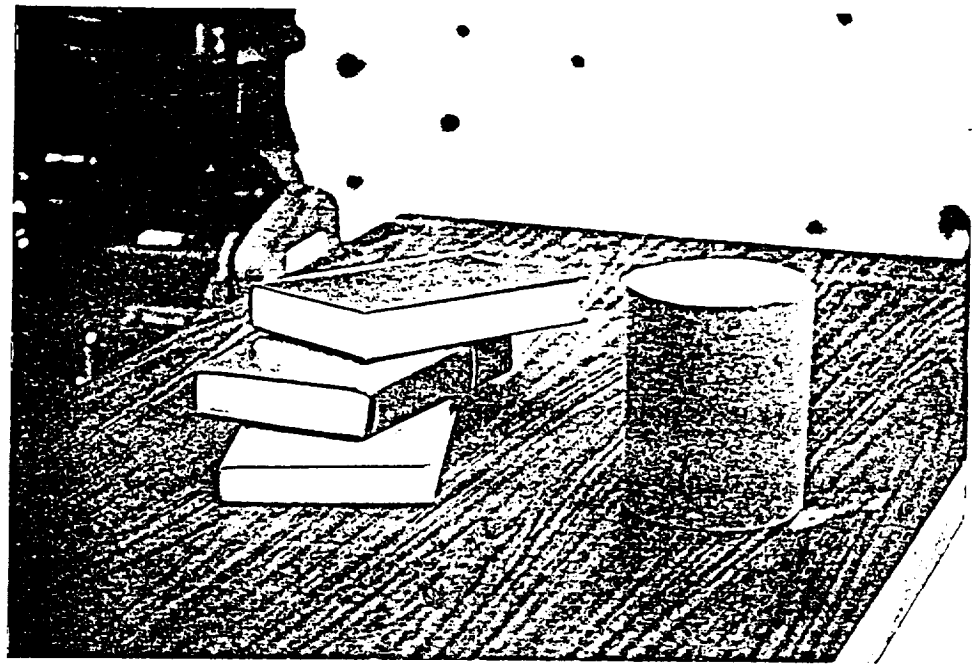


Figure A.3: Books and cylinder used during surface map application

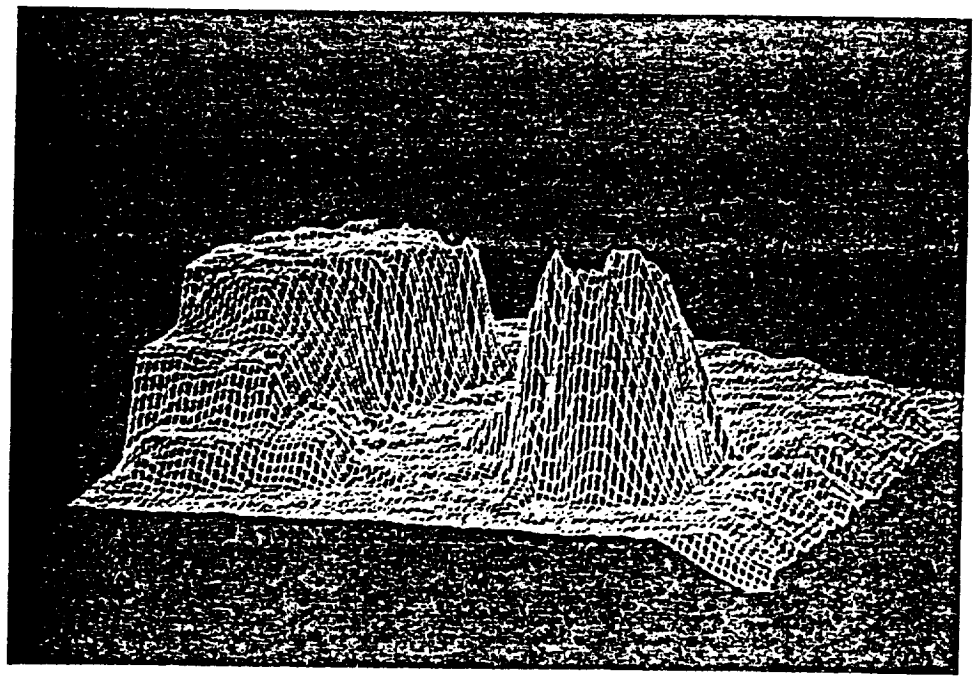


Figure A.4: Range map of books and cylinder convolved with a 3 × 3 low pass filter

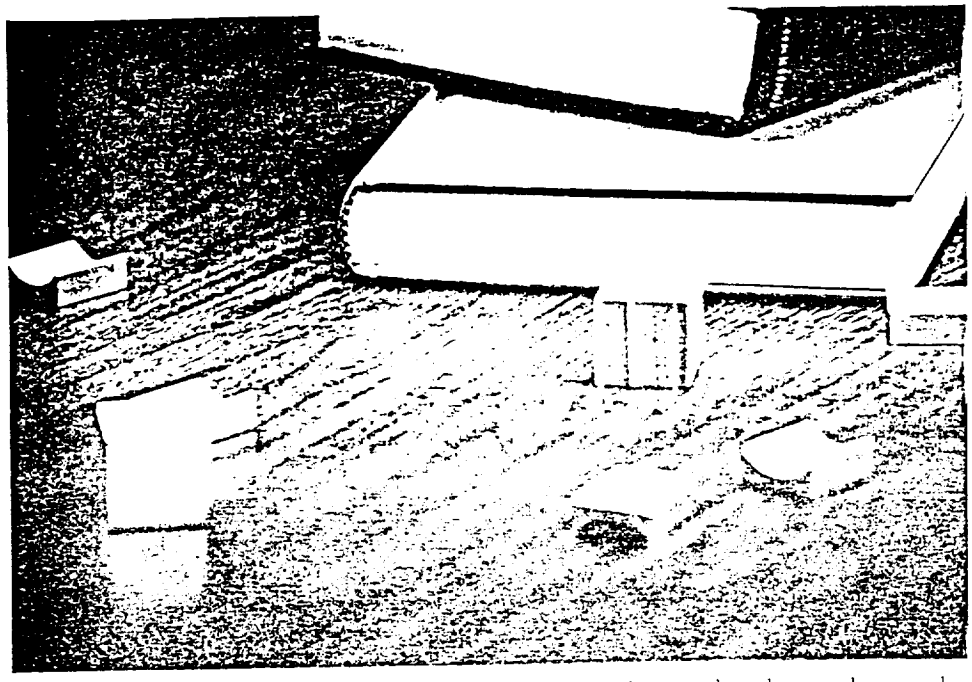


Figure A.5: Scenario of laser spot in close proximity to books and metal objects. Maximum field of view of picture is $\approx$ 30cm

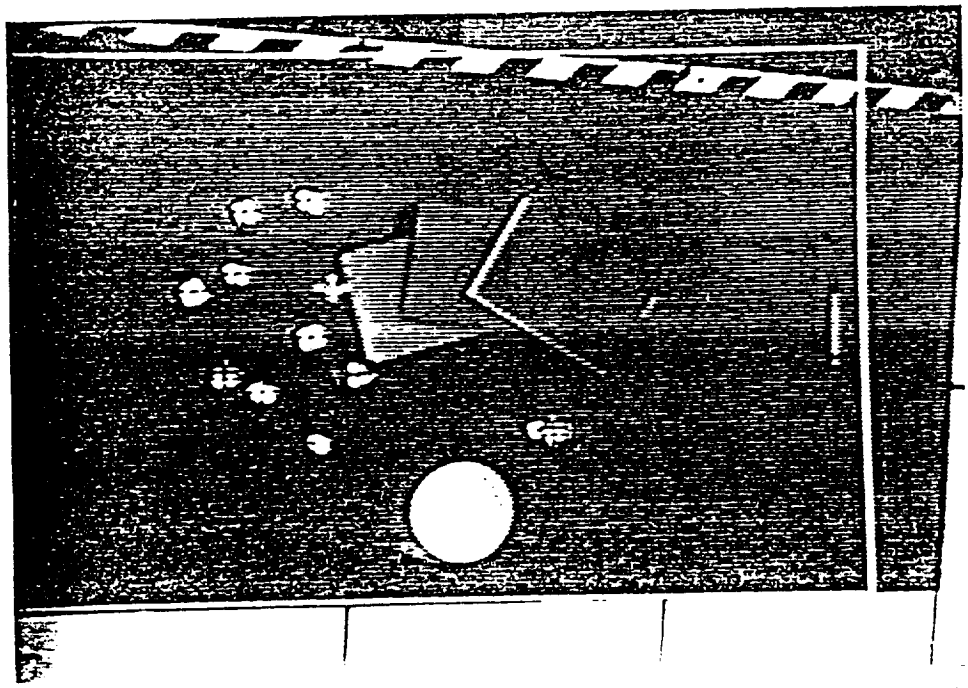


Figure A.6: View of the same scenario from within a computer. Specular reflections caused by metal objects are identified with blue crosses as having passed the size and intensity constraints. Red cross is the laser spot which is [illegible] remaining region remains [illegible]